ENBRIDGE

Andrew Mandyam
Director Regulatory Affairs
& Financial Performance

tel 416-495-5499
fax 416-495-6072
EGDRegulatoryProceedings@enbridge.com

Enbridge Gas Distribution
500 Consumers Road
North York, Ontario M2J 1P8
Canada

December 19, 2017

VIA RESS, EMAIL and COURIER

Kirsten Walli
Board Secretary
Ontario Energy Board
2300 Yonge Street, Suite 2700
Toronto, Ontario, M4P 1E4

Dear Ms Walli:

Re: Enbridge Gas Distribution Inc. ("Enbridge")
Ontario Energy Board ("Board") File Number EB-2017-0324
Application for 2015 Demand Side Management ("DSM") Clearance of Deferral and Variance Accounts

Enclosed is the application and supporting evidence of Enbridge concerning the final disposition and recovery of the 2015 DSM deferral and variance accounts and the request for approval for disposition of these amounts at the next available QRAM application following a Decision of the Board.

Please contact the undersigned if you have any questions concerning this application and evidence.

Yours truly,

(Original Signed)

Andrew Mandyam
Director, Regulatory Affairs & Financial Performance

Enclosures
cc: Dennis O'Leary, Aird & Berlis LLP

EXHIBIT LIST

EXHIBIT A - ADMINISTRATION

EXHIBIT	TAB	SCHEDULE	DESCRIPTION
A	1	1	Exhibit List
	1	2	Application
	1	3	Application for Clearance of DSM Accounts
			Request for Approval and Clearance of 2015 DSM Deferral and Variance Accounts

EXHIBIT B – EVIDENCE

EXHIBIT	TAB	SCHEDULE	DESCRIPTION
B	1	1	2015 DSM Final Annual Report, December 18, 2017
	2	1	2015 DSM Deferral and Variance Account Rate Allocation and Clearance of 2015 DSM Balances
		2	Estimated Typical Bill Impacts based on Rate Allocation
	3	1	Enbridge Responses to 2015 Annual Verification Recommendations made by the Evaluation Contractor
	4	1	OEB Letter of Direction RE: 2015-2020 Demand Side Management Evaluation Process of Program Results EB-2015-0245, dated August 21, 2015

ONTARIO ENERGY BOARD

IN THE MATTER OF the Ontario Energy Board Act, 1998, S.O. 1998, c. 15, Schedule. B, as amended;

AND IN THE MATTER OF an application by Enbridge Gas Distribution Inc. for an order or orders approving the balances and the clearance of certain Demand Side Management Variance Accounts into rates, within the next available QRAM following the Board's approval.

APPLICATION

1. Enbridge Gas Distribution Inc. ("Enbridge" or the "Company") is an Ontario corporation with its head office in the City of Toronto. It carries on the business of selling, distributing, transmitting and storing natural gas within Ontario. The Company also undertakes Demand Side Management ("DSM") activities.

2. Enbridge hereby applies to the Ontario Energy Board (the "OEB" or the "Board"), pursuant to section 36 of the Ontario Energy Board Act, 1998, as amended (the "Act"), for an Order or Orders approving the final balances in the following 2015 DSM deferral accounts and the disposition of these balances:

	$	10,077,695
Demand Side Management Variance Account (DSMVA)	$	825,460
Lost Revenue Adjustment Mechanism Variance Account (LRAMVA) (Reimbursable to Ratepayers)	$	(71,829)
Total Amount Recoverable	$	10,831,326

3. Enbridge applies to the Board for such final and interim orders and/or accounting orders as may be necessary in relation to the clearance of the accounts which are

the subject of this Application, within the next available QRAM following the Board's approval. The Company further applies to the Board pursuant to the provisions of the Act and the Board's Rules of Practice and Procedure for such final and interim Orders and directions as may be necessary in relation to this Application and the proper conduct of this proceeding.

4. The persons affected by this Application are the customers of Enbridge. It is impractical to set out the names and addresses of the customers because they are too numerous.

5. Enbridge requests that a copy of all documents filed with the Board by each party to this proceeding be served on the Applicant and the Applicant's counsel, as follows:

Mr. Andrew Mandyam
Director, Regulatory Affairs
Enbridge Gas Distribution Inc.

Address for personal service:	500 Consumers Road North York, ON M2J 1P8
Mailing Address:	P.O. Box 650 Scarborough, ON M1K 5E3
Telephone:	416.495-5499
Facsimile:	416.495-6072
E-mail:	EGDRegulatoryProceedings@enbridge.com

Please quote the name or docket number of the proceeding in all communications.

The Applicant's counsel:

Mr. Dennis M. O'Leary
Aird & Berlis LLP

Address for personal service and Mailing address:	Brookfield Place, Box 754 Suite 1800, 181 Bay Street Toronto, ON M5J 2T9
Telephone:	416-865-4711
Facsimile:	416-863-1515
E-mail:	doleary@airdberlis.com

Dated: December 19, 2017, Toronto, Ontario.

ENBRIDGE GAS DISTRIBUTION INC.

Per:_____(Original Signed)___________

APPLICATION FOR CLEARANCE OF DSM ACCOUNTS

Request for Approval and Clearance of 2015 DSM Deferral and Variance Accounts

1. Enbridge Gas Distribution Inc. ("Enbridge" or the "Company") is applying to the Ontario Energy Board ("OEB" or the "Board") pursuant to Section 36 of the *Ontario Energy Board Act, 1998,* as amended ("Act") for an Order or Orders approving the final balances in certain 2015 Demand Side Management ("DSM") Deferral and Variance Accounts. The Company is also seeking approval for the disposition of the balances in these accounts through a one-time adjustment in rates, within the next available QRAM following the Board's approval.

2. As outlined in the Filing Guidelines to the DSM Framework for Natural Gas Distributors (2015 to 2020) (EB-2014-0134) ("Guidelines"): "Consistent with past practices, recovery and disposition of DSM related amounts (i.e., DSM Variance Account ("DSMVA"), DSM Incentive Deferral Account ("DSMIDA"), and LRAM Variance Account ("LRAMVA")) will be filed by the natural gas utilities annually, based on the actual amount of natural gas savings resulting from the utilities' DSM programs in relation to the annual plans targets. The DSM amounts include program spending, shareholder incentive amounts and lost revenues in relation to the DSM programs delivered by the natural gas utility."[1]

3. The deferral and variance accounts which are the subject of this proceeding relate to DSM activities in 2015. Though the current Framework encompasses 2015 to 2020, the Board directed that 2015 would act as a transition year and the "gas utilities should roll-forward their 2014 DSM plans, including all programs and

[1] Filing Guidelines to the DSM Framework for Natural Gas Distributors (2015-2020) (EB-2014-0134), page 36

parameters (i.e., budget, targets, incentive structure) into 2015."[2] The accounts which are the subject of this Application and the balances recorded are as follows:

Table 1.

2015 DSM Deferral and Variance Accounts and Balances	
Demand Side Management Variance Account (DSMVA)	$ 825,460
Demand Side Management Incentive Deferral Account (DSMIDA)	$ 10,077,695
Lost Revenue Adjustment Mechanism Variance Account (LRAMVA) (Reimbursable to Ratepayers)	$ (71,829)
Total Amount Recoverable	$ 10,831,326

4. New to the evaluation and audit process for the 2015 DSM year results, as directed by the Board, was the transfer of responsibility for oversight of the annual evaluation, measurement and verification ("EM&V") process from the utilities to OEB Staff. For reasons set out in the evidence below which detail Enbridge's concerns about the EM&V process which was followed and the results, this application reflects all 2015 verified program results as presented in the Evaluation Contractor's ("EC") final report: *The Ontario Gas DSM Evaluation Contractor 2015 Natural Gas Demand Side Management Annual Verification report* by DNV GL ("DNV" or the "EC") dated October 12, 2017,[3] ("the EC Report") with the exception of the incomplete Net-to-Gross ("NTG") Study findings (encompassing Custom Commercial, Custom Industrial and Run It Right offers).[4]

[2] EB-2014-0134, Report of the Board, DSM Framework for Natural Gas Distributors (2015-2020), Sect. 4.2, page 30
[3] Filed in EB-2015-0245
[4] In the process of preparing this Application, Enbridge determined there were a number of errors made by the EC in its calculation of verified 2015 DSM program results including its determination of DSM shareholder incentive and LRAM. These errors were also reflected in the audit opinion provided in the EC Report date October 12, 2017. Enbridge communicated these concerns in an email to Board Staff and the EC on November 20, 2017. At a meeting of the EAC on December 6, 2017, the EC acknowledged errors in their calculations. In an email to the EAC

5. More specifically, Enbridge has identified the following primary concerns with the OEB Staff coordinated 2015 verification process and results:

- The evaluation and audit process lacked the appropriate and necessary degree of transparency, collaboration, efficiency and balanced stakeholder input to ensure a fair and credible process and result;

- The retroactive application of the NTG ratios from the NTG Study is inappropriate and contrary to the Board's earlier Direction and is both inappropriate and inconsistent with best practices.

- The determination of NTG ratios in the NTG Study by DNV are inappropriate and flawed in that the NTG Study deviated from the appropriate scope of work and did not reflect industry best practice.

6. Each of the concerns listed above are described and explained within the body of this evidence. For the purposes of this application, Enbridge is applying for approval by the Board of the amounts listed in Table 1. These are the product of all 2015 verification results recommended by the EC with the exception of the flawed and inappropriate retroactive application of incomplete NTG Study results.

7. As outlined in the Guidelines: "The Board expects that the utilities will use the results of the Final Audit & Evaluation Report when they file for disposition of their respective DSM deferral and variance accounts."[5] Section 11.0 of the Guidelines states that:[6]

---

on December 13, 2017, the EC outlined corrected calculations and Enbridge expects that these corrected values will be included by the EC in updated final reports. At the time of filing this application, no updated reports from the EC have been issued by the OEB, however the EC's corrected values have been incorporated into the determination of the Deferral and Variance Account balances outlined by Enbridge in this application.

[5] Filing Guidelines to the DSM Framework for Natural Gas Distributors (2015-2020) (EB-2014-0134), page 20

[6] Ibid., page 37

> The natural gas utilities should apply annually for the disposition of any balances in their LRAMVA and DSMVA and, if applicable, apply for a shareholder incentive amount associated with the previous DSM program year and disposition of any resulting DSMIDA balance.
>
> This application should include the final results as outlined in the Final Evaluation and Audit Reports, and information setting out the allocation across rate classes of the balances in the LRAMVA, DSMVA and DSMIDA.

In accordance with Section 11.0 of the Guidelines and for comparative purposes, Enbridge provides the DSM values based on the EC's recommended amounts and Enbridge's application as seen in Table 2 below:

Table 2.

2015 DSM Achieved Savings, Shareholder Incentive, and Lost Revenue	Enbridge Pre-Audit	Audit Opinion of EC*	Enbridge Application
Shareholder Incentive	$10,318,594	$ 6,207,339	$ 10,077,695
Lost Revenue	$28,800	$ 16,405**	$28,976
DSMVA	$825,460	$825,460 (not reviewed)	$825,460

* As noted previously in footnote 4, Enbridge determined there were a number of errors made by the EC in its calculation of verified 2015 DSM program results. The Shareholder Incentive originally recommended in the EC's Final Verification report dated October 12, 2017 was $6,489,467. As a result of Enbridge identifying errors in the manner the EC had applied its verification adjustments to the program results, the EC undertook a review and provided revised figures (as outlined above) to the EAC on December 13, 2017. The EC has advised it will be updating the final reports originally dated October 12, 2017 to reflect these corrected values.

** The above table includes the EC calculated and the Enbridge calculated Lost Revenue values for comparison to align with audit opinion category presented in the EC's Annual Verification Report. For clarity, the LRAMVA value requested for disposition in this application is outlined in Table 1.

8. Notwithstanding the issuance of the EC Report, there are several reasons why it should not be accepted and relied upon for a final determination of amounts for Clearance. First, it undertook its calculations employing an approach, in Enbridge's determination, which is contrary to Board decisions and policy – discussed further below.

9. Second, it excludes an important required feature outlined in the scope of work of the NTG Study, namely Enbridge/Union Gas program based determinations of spillover. While it does include a proxy deemed spillover value sourced from another study conducted in Massachusetts (applied as a result of an instruction given by Board Staff – to be discussed further below), Enbridge views the report as incomplete.

10. Third, the EC Report excludes another important feature of the NTG Study specified in the scope of work, namely Secondary Attribution. DNV quantified Secondary Attribution but did not apply these findings to final NTG Study results.

11. Fourth, the NTG outcomes are not credible and Enbridge does not have confidence in them as they do not reflect best practice approaches in undertaking self-report NTG studies.

12. Fifth, Enbridge uncovered a number of errors made by the EC throughout the verification process, including in the evaluator's individual project savings verifications, as well as in the EC's application of adjustments to arrive at verified 2015 DSM program results; most importantly errors were made in the EC's final determination of verified net cumulative savings and subsequently, the DSM shareholder incentive and LRAM. These errors primarily related to how the EC applied the CPSV verification adjustments across the total custom project results.

13. Finally, in Enbridge's efforts to gain understanding of NTG adjustments made by the EC, despite continued requests for detailed information to enable the Company to replicate the calculations used by the EC to arrive at its proposed NTG values, the EC failed to provide the details required for the Company to do this analysis. Enbridge therefore had no ability to review live calculations or understand the consideration of participant responses to the NTG scoring algorithm. Given errors already uncovered by Enbridge, with great effort, in other areas of the verification

where the Company was successful in obtaining data, Enbridge does not have confidence in the results and has no way to assess how the NTG calculations were done, if they were done correctly or what may need to be reviewed. This "black box" determination adds to Company's uncertainty of the results.

14. For these reasons, discussed in further detail below, the utility is of the view that the Board should not have confidence in the determination of the NTG Study values.

15. This evidence has been organized based on the following general outline. Section 1 describes the 2015 EM&V process, highlights the significant delay, and concludes that a number of issues and events demonstrate a lack of transparency and create concern with objectivity within the current process that was prejudicial to the Company. Section 2 addresses the issue of retroactivity and the inappropriateness for inclusion of the EC's NTG Study recommendations in the determination of 2015 program outcomes. Section 3 details many of the concerns with the NTG Study undertaken by DNV that cause the Company to conclude that the results are unreliable and should not be accepted in a number of material respects. Finally, Section 4 outlines the summary outcomes as a result, and forms the basis of the application by Enbridge made herein.

Section 1 – The 2015 EM&V Process

<u>The 2015 DSM Evaluation Process – Summary of Facts and Events</u>

16. The Board issued the Report of the Board, DSM Framework for Natural Gas Distributors (2015-2020) (EB-2014-0134) ("Framework") and the Guidelines on December 22, 2014. Previously, as directed by the Board, the evaluation process relating to DSM programs had been a function that the gas utilities managed with input from stakeholders throughout the process. The prior approach was supported

by the Technical Evaluation Committee ("TEC") and the Audit Committee which included stakeholder representatives chosen by the stakeholder community.

17. In this framework, the TEC established DSM technical and evaluation standards for the natural gas utilities in Ontario. The TEC consisted of seven individuals: three intervenor members, a representative from Union, a representative from Enbridge, and two independent members with technical and other relevant expertise. The Audit Committee for each utility consisted of three intervenor members and one utility representative. In the 2015-2020 Framework however, the Board concluded that it was "in the best position to coordinate the evaluation process throughout the DSM framework period"[7] in collaboration with the gas utilities, supported by stakeholders with technical expertise. The Guidelines further specified that "the Board will take on the coordination function of the EM&V process."[8]

18. The Board subsequently issued two letters on August 21, 2015 and March 4, 2016 which further outlined the new evaluation process and the transition of the activities of the TEC to the OEB.

<u>August 21, 2015 Letter from the Board</u> [9]

19. The Board's August 21, 2015 letter outlined a new DSM evaluation and audit governance structure. Specifically, this letter outlined the DSM evaluation governance structure, the evaluation approach, and the roles of the parties involved in the evaluation process. The letter also introduced the formation of an Evaluation Advisory Committee ("EAC") and explained that once the OEB had retained an EC, OEB Staff would work with the TEC to transition work already

[7] EB-2014-0134, Report of the Board, DSM Framework for Natural Gas Distributors (2015-2020), Sect. 4.2, page 30
[8] Filing Guidelines to the DSM Framework for Natural Gas Distributors (2015-2020) (EB-2014-0134), page 15
[9] Letter from the Board, 2015-2020 Demand Side Management Evaluation Process of Program Results, August 21, 2015 (EB-2015-0245)

ongoing under the responsibility of the TEC to the EAC. In the meantime, the OEB directed the gas utilities and the TEC to continue working on the evaluation projects that they had initiated until the transition occurred.

20. As stated in the Board's August 21, 2015 letter, the EAC was to provide input and advice throughout the process, including the evaluation and audit of DSM results and the development of the Evaluation, Measurement & Verification Plan to be drafted by the EC.

21. The letter outlined that the EAC would be comprised of:
    - Experts representing non-utility stakeholders, with demonstrated experience and expertise in the evaluation of DSM technologies and programs
    - Expert(s) retained by the OEB
    - Representatives from the IESO
    - Representatives from each natural gas utility
    - Representatives from the Ministry of Energy (MOE) and the Environmental Commissioner of Ontario (ECO), who will participate as observers.

22. The Board's August 21, 2015 letter announced the appointment of the following non-utility stakeholders to the EAC: Marion Fraser, Marion Fraser Enterprises Inc.; Chris Neme, Energy Futures Group; and, Jay Shepherd, Shepherd Rubenstein Professional Corporation.

March 4, 2016 Letter from the Board [10]

23. The Board's letter dated March 4, 2016 outlined the transition of the, then current, ongoing DSM evaluation activities from the TEC to the EAC. Before the formation of the EAC and the hiring of an EC, a number of important evaluation activities were already underway. Among several key projects, the Custom Commercial and Industrial NTG Study was in progress. The TEC had previously completed a Request for Proposal ("RFP"), initiated a selection process, and had contracted DNV (previously DNV Kema) in May 2015 to complete the study. As acknowledged in DNV's scope of work (see Exhibit B, Tab 5, Schedule 2), the TEC had resolved in April, 2014 that "the primary objective of this project is a transparent, reputable study that produces strong, credible, and defensible NTG ratios to be used on a go-forward basis"[11] (emphasis added).

24. At the time the Board issued its March 4th letter, the TEC had done considerable work with DNV to refine the details of the study and DNV had also drafted an updated Scope of Work (see Exhibit B, Tab 5, Schedule 2) to the TEC outlining its approach. The objective for the study stated that the "goal of this evaluation is to develop transparent free ridership and spillover factors for custom commercial and industrial programs, to be used for future programs"[12] (emphasis added).

25. The March 4, 2016 letter acknowledged that the NTG Study work plan was to be presented to the TEC at its next meeting scheduled for March 10, 2016 (Board Staff was in attendance at this meeting). The letter further provided that "following input from the TEC, this study will be transitioned to OEB. The utilities will continue to manage contractual obligations and payments associated with this project. OEB

[10] Letter from the Board, Transition of Technical Evaluation Committee Activities to the OEB, March 4, 2016 (EB-2015-0245)
[11] Measurement of NTG Factors for Ontario's Natural Gas Custom Commercial and Industrial DSM Scope of Work for Ontario Natural Gas Technical Evaluation Committee (TEC), dated March 2, 2016, page 7
[12] Ibid., page 4

Staff will assume oversight of the study and will confirm the completion of major milestones for the utilities to process payments of consultant's invoices."[13]

26. Importantly, the March 4, 2016 letter from the Board did not contemplate that OEB Staff could unilaterally alter or change the scope of DNV's work already underway.

Board Staff Coordinated 2015 EM&V Process

27. On May 5, 2016, an email was sent to members of the EAC from OEB Staff. The email announced that the Board had appointed Dr. Ted Kesik, University of Toronto and Bob Wirtshafter, Wirtshafter Associates, Inc. as independent experts on the EAC. The email acknowledged that Mr. Wirtshafter and Mr. Kesik had served as members of the former TEC that assisted in providing objective advice during the previous DSM Framework.

28. In April 2016, the Board selected the EC. The May 5, 2016 email outlined that OEB Staff had engaged DNV as the EC. It indicated that among the EC's responsibilities, DNV would oversee the annual verification of the 2015 DSM program results, including preparing a Final DSM Results Report. The EC was selected with no EAC or utility input or consultation. As outlined above, DNV had already been selected by the TEC the year prior to complete the custom commercial and industrial NTG Study as previously detailed.

29. The 2015 EM&V process took approximately 18 months (it should be noted however that in this timeframe the NTG study was not completed, as the spillover component is as yet incomplete). In the first year of the new OEB Staff led EM&V process, almost 22 months after the end of the utilities' 2015 program year, the OEB issued two reports on October 16th, 2017, developed by the EC, DNV, providing its calculations for 2015 DSM verification results.

[13] Letter from the Board, Transition of Technical Evaluation Committee Activities to the OEB, March 4, 2016 (EB-2015-0245)

30. It should be noted that the evaluation process envisioned by the Board was a "process coordinated by the Board, in collaboration with the gas utilities, and supported by stakeholders with technical expertise, will be one that results in a thorough evaluation of DSM programs in an efficient manner. By taking on a larger role in the EM&V process, the Board will consult and seek expert opinion from both the gas utilities and stakeholders as appropriate"[14] Unfortunately, the 2015 EM&V process did not exhibit the collaboration, transparency and efficiency intended by the Board which has led to much uncertainty, instability, and faulty outcomes.

Lack of Transparency and Collaboration within the Process

31. As outlined in the Guidelines, at a minimum the Board expects the independent third party auditor will be asked to[15]:
    - Review the draft evaluation reports prepared by the gas utilities and verify the components of the draft program results;
    - Conduct audits of DSM programs to ensure that the results proposed by the gas utilities are accurate;
    - Confirm the calculations of savings and the draft evaluations conducted by the gas utilities are consistent with the evaluation plans approved by the Board;
    - Provide an audit opinion on the DSMVA, lost revenues and shareholder incentive amounts proposed by the natural gas utilities and any subsequent amendments;
    - Confirm any target adjustments have been correctly calculated and applied;
    - Identify any input assumptions that either warrant further research or that should be updated with new best available information;
    - Review the reasonableness of any verification work that has been undertaken by the gas utilities and included in the Draft Evaluation Reports;

[14] EB-2014-0134, Report of the Board, DSM Framework for Natural Gas Distributors (2015-2020), Page 30
[15] Filing Guidelines to the DSM Framework for Natural Gas Distributors (2015-2020) (EB-2014-0134), page 19

- Recommend any forward-looking evaluation work to be considered; and,
- Prepare a Final Audit & Evaluation Report.

32. Though the EC is expected to act on this mandate, the lack of transparency and at times poor communication through overseeing the 2015 process, specifically the Custom Project Savings Verification process as well as the NTG Study leaves Enbridge with the belief that Board Staff chose to direct, rather than oversee, the EC without the benefit of Enbridge's experience and expertise in DSM, or more broadly without full consideration of all information and expertise provided through the EAC. Enbridge is of the view this is clearly contrary to the intention of the Board's specific articulation that Board Staff would be "coordinators", and therefore that the EC would remain an independent, third party auditor.

33. Of particular concern to Enbridge in the 2015 process were two significant examples where OEB Staff explicitly instructed the EC on how to proceed with: i) the application of the NTG study, and ii) the finalization of the spillover component of the NTG results.

34. As further discussed below, the outcomes of the NTG Study should not be applied to 2015 program results for the purpose of determining the Company's shareholder incentive. However, following the transition of the oversight of the NTG Study from the TEC to Board Staff in March of 2016, Enbridge requested clarity and understanding of Board Staff's proposal regarding the application of NTG Study outcomes to 2015 program results. Ultimately, a year later, the day prior to the EAC receiving a copy of the EC's draft CPSV/NTG report, Board Staff emailed the two utilities on May 23, 2017, and confirmed it had instructed DNV to retroactively apply the NTG Study results (they were not in fact NTG values, they proposed free ridership values but did not include spillover) to 2015 DSM program results. Board Staff indicated this was in line with their understanding of the Board's direction.

This action was neither appropriate in light of Board Staff's role as coordinator in the EM&V process; nor correct with respect to the Board's Decision and Order of January 20, 2016 and revised Decision and Order of February 24, 2016.

35. Further, during a conference call with the EAC on September 27, 2017, Board Staff communicated that the EC had been instructed to find a proxy deemed spillover value to be applied to the utilities' 2015 program results, notwithstanding the fact that the spillover study was still ongoing and incomplete (though when queried by various members of the EAC, it was unclear whether these instructions were provided to the EC by OEB Staff or the OEB). Board Staff indicated that, rather than wait for the final results of the EC's spillover research, the EC was directed to instead find and apply a deemed value to approximate spillover effects based on a spillover value in another jurisdiction and proceed to finalize the 2015 program results verification. Enbridge is of the view that Board Staff's priority at this point was simply to rush to complete the 2015 EM&V process, given the significant amount of time that had already transpired. Enbridge submits that taking such action is outside of Board Staff's role as coordinator of the EM&V process and consisted of a unilateral change to the scope of work outlined for the NTG Study.

36. On the issue of consultation with stakeholders, direction was provided by the Board in its 2015-2020 Framework for Demand Side Management. The Board clearly acknowledges the utilities' expertise and experience with regards to DSM:

> Although the Board's role will be increased, primarily with respect to oversight related to the evaluation process and annual updates to the input assumptions list, the Board continues to see the direct involvement of all key stakeholders, notably the gas utilities and intervenors with the required expertise, to be critical and necessary to ensure all elements of the gas utilities' multi-year DSM plans are considered during the program development, approval and evaluation stages.[16]

[16] EB-2014-0134, Report of the Board, DSM Framework for Natural Gas Distributors (2015-2020), Page 36

37. In the Framework, the Board concluded that it was “in the best position to coordinate the evaluation process throughout the DSM framework period”[17] in collaboration with the gas utilities, supported by stakeholders with technical expertise. The Guidelines further specified that “the Board will take on the coordination function of the EM&V process.”[18], however, no clear definition of OEB Staff’s role as “coordinator” was provided, nor was there clear direction given regarding the relative roles and responsibilities of the EAC in regard to the decision making process, particularly with regard to input into the selection of evaluation experts, managing stakeholder input to evaluation scopes, and resolution of diverging viewpoints and conflicts of interest. The utilities were consistent in their request for clarity in this regard, including the repeated recommendation for the development of an EAC charter, clear project management oversight, the provision of detailed and maintained timeline plans, as well as the recording and distribution of EAC minutes.

38. Effectively, the Company believes the EAC’s structure and the burden placed on OEB Staff, with limited prior experience, contributed to a lack of transparency and limited collaboration between Board Staff and members of the EAC. As a result, the Company urges the Board to direct the Evaluation Advisory Committee to create a consensus-based charter that encourages cooperation between all parties, so that the expertise and experience from each party may be applied to the evaluative challenges inherent in the process. We understand Board Staff is currently drafting a charter in an effort to establish greater role clarity for the EAC.

39. Enbridge notes that it developed together with Union Gas and other DSM stakeholders a document entitled Joint Terms of Reference which was filed in draft and ultimately approved by the Board in EB-2011-0295. This document detailed

[17] Ibid., Section 4.2, page 30
[18] Filing Guidelines to the DSM Framework for Natural Gas Distributors (2015-2020) (EB-2014-0134), page 15

the duties and responsibilities of all participants in the EM&V process during the 2012 to 2014 Framework and was of great assistance ensuring that the review of annual results and the updating of measure assumptions was undertaken in an objective and efficient fashion. These rules guided the parties and provided certainty as to the process. Enbridge submits that a charter which includes a materially similar set of rules would greatly assist in the timely generation of credible results in future.

40. Enbridge has reviewed and provided input in the development of the draft charter which Union Gas has prepared and appended to its 2015 DSM Clearance Application and supports the Board's adoption and approval of this charter for immediate implementation. By broadening the decision making process as contemplated in this charter, all parties will be inclined to take ownership of the process and improve the quality of the outcome of the EAC.

41. With regard to one of the most significant evaluation efforts, the development of the NTG Study, though the EC did solicit comments from the EAC on the survey instrument, much of the commentary and input provided by Enbridge was dismissed. In addition Enbridge was not provided an opportunity to provide input in respect of the EC's determination of an appropriate scoring algorithm and upon receipt of the draft results of the free-ridership interviews, Enbridge repeatedly requested details regarding the determination of participant scoring based on example feedback provided, but has received limited information from this request.

42. In addition, despite repeated requests by Enbridge for the EC to provide complete details of the data used in its determinations, in many regards, the EC would not provide Enbridge with detailed documentation or clear calculations to allow Enbridge to replicate (and therefore understand and confirm) the EC's findings. Given that there were instances where Enbridge was able to work through the data

to verify the EC's calculations and errors were found, the lack of detail and transparency in other regards, leaves Enbridge uncertain of the determinations of the EC. Enbridge is of the view that there should be full transparency in the process to allow the Company the opportunity to fully review adjustments.

43. In October 2017, the US Department of Energy updated its Uniform Methods Protocols ("UMP") on Self-Report surveys. Enbridge has included a copy of the Uniform Methods Protocols at Exhibit B, Tab 6, Schedule 4. The UMP is widely regarded as the industry standard for how to conduct evaluative surveys. Of note in the recent update is a detailed commentary on the critical importance of transparency between surveyors and interested parties. "Ensuring transparency" is identified as one of the 6 key principles of best practice. Citing numerous recent U.S. studies, the UMP stresses "the importance of making the entire process transparent so stakeholders can understand how each question and its response impacts the final estimate."

44. The UMP delves into significant detail on the involved role stakeholders should play in the development and execution of Net-to-Gross surveys, explicitly recommending that "jurisdictions should design evaluation plans to assess net savings in conjunction with the key stakeholders". Survey components to be shared with stakeholders include "details of critical elements such as the question sequence, scoring algorithms, and the handling of inconsistent and/or missing data."

45. Enbridge is concerned that the process undertaken for the 2015 EM&V process lacked sufficient consultation and collaboration to generate the most effective, fair, and reliable results. In comparison to prior experiences with third party evaluators, Enbridge believes the process underpinning the 2015 EM&V results, in particular

the NTG study, was less transparent and less certain, and therefore less credible, despite an unfortunate over expenditure in terms of both time and cost.

46. A concern for Enbridge involved the practice of OEB Staff receiving and reviewing reports and deliverables from the EC prior to the EAC. During the verification process, it became clear that OEB Staff were providing comments and feedback to the EC that were not visible to the EAC, including undisclosed comments on specific evaluation reports. The utilities requested transparency in this regard; however comments were not shared with the group. Though the EC's incorporation of, or impact of these comments are not known, this lack of transparency caused concerns regarding the ability of the EC to maintain the position of an independent expert and brought the objectivity of Board Staff's role as overseer into question.

47. Other examples that contributed to a less than transparent and credible process included:

- a refusal to record meeting minutes to capture key decision (despite suggestions from the utilities to do so);
- failing to track and follow up on meeting action items;
- questions and decision points that went unanswered creating uncertainty; and,
- a refusal to provide clarity and transparency regarding budgets and spending for the EM&V related activities.

Despite the utilities having responsibility and accountability for an overall annual evaluation budget for their respective DSM portfolios, OEB Staff has refused to provide details on EM&V budgets for planned verifications or details regarding forecasted spending in a given year. At a minimum, the utilities require budget information to facilitate contract payment, to assist with budgeting for other aspect of program planning and utility led evaluation (e.g. process evaluation) as well as to support financial reporting requirements. Currently Enbridge has no ability to monitor spending or accrue funds, this has proved unnecessarily challenging.

Delays in the 2015 EM&V Process and Impacts to Enbridge

48. In previous years, in consultation with the Audit Committee, Enbridge aimed to contract an auditor for the current program year in October of the same year to enable the process to be completed by June 30th of the year following to meet the Reporting & Record Keeping Requirements Rule for Gas Utilities. For the 2015 program year, Board Staff issued an RFP in early 2016 to facilitate the selection of the EC which specified the completion of a Final Results Report in October 2016. However, following the selection of the EC in April 2016, there were significant delays throughout. The kick-off meeting of the EAC and the EC was on May 12, 2016. The EC's draft of an overall EM&V plan was not provided to the EAC until September 2016 and was not finalized until February 2017. The work plan for the CPSV/NTG evaluation was not finalized until December 2016. Though Enbridge had provided its custom project tracking data to DNV in the early part of 2016 when the NTG project was previously underway at the TEC, after Board Staff assumed the oversight of the NTG study and changed the scope of work to a combined CPSV/NTG effort, the CPSV verification and the NTG Study did not begin execution until late January 2017. These delayed timelines meant that the EC did not distribute a draft verification report until late July 2017 and ultimately, the OEB issued the EC's final reports on October 16th, 2017 (without having completed the spillover component of the NTG study). This was one full year after the date contemplated in the Board issued RFP for an Evaluation Contractor, and almost two years after the end of the 2015 program year.

49. Though in transitioning to the new EM&V process, it could be reasonably expected that there might be some delay, Enbridge would not have anticipated this outcome.

In particular, Enbridge shared concerns about how delayed evaluation efforts impacted and inconvenienced customers who were being queried on projects that were implemented over a year, and in some cases, over two years previous. This impacted the ability for the EC to connect with customer contacts that had sufficient (or any) knowledge of specific projects and most certainly impacted customers' recall regarding projects details and arguably effected NTG responses.

50. It is important also to note that while the EC and OEB staff often missed timelines and deadlines, the utilities were given very little opportunity to provide input to project timelines and were regardless expected to meet aggressive deadlines dictated to them, in providing data or responses. For example, Enbridge was required on multiple occasions to rush its review and comments on very large quantities of CPSV data due to short, unrealistic deadlines set by the EC, working under the direction of Board Staff, and on occasion despite not being provided all of the information necessary to complete the requested review. In a number of instances where Enbridge felt it necessary to raise concerns with OEB Staff, no response or feedback was provided and due consideration was often not afforded.

51. The point here is not to be overly critical of Board Staff's efforts. They simply did not have the experience and capacity to fully manage the undertaking and there was obvious pressure on both the EC and the utilities to meet aggressive timelines, to make up for significant early delays in Board Staff's initiation of the 2015 process. The result is therefore not likely what the Board had envisioned at the outset.

Change in NTG Study Scope

52. As described above, the TEC had originally scoped the study of work to be done with regard to the NTG Study. The study was meant to, "provide guidance on the development of a strategy for applying free ridership and spillover data collected on previous program participants to forward looking DSM program activity[19]" [emphasis added]. It was never contemplated that the results would be applied retroactively.

53. Working within a consensus-based decision-making process, the TEC had resolved that the prospective application was the most appropriate way forward with regards to NTG study findings. The prospective approach was also endorsed by the contracting consultant (DNV) as appropriate. The sample design originally proposed by DNV was one that was well suited for prospective, forward looking application of results. Only after Board Staff became involved in overseeing the NTG Study did this component of the study undergo a critical change, despite concerns raised at the EAC on multiple occasions by the utilities. This began with a new proposal with regard to the sample design and methodology to be employed; DNV presented the new approach to overlap with the CPSV verification for application retrospectively and specifically to the population of 2015 project results.

54. In the Company's view, Board Staff altered the scope of work for the EC in this regard, directing it to misapply Board policy. While Enbridge comments on the policy issue of retroactivity further in Section 2 of this evidence, it is appropriate to consider here the role of Board Staff in an objective and transparent EM&V process. The Company submits that Board Staff's role in the evaluation and audit

[19] Ontario Natural Gas Technical Evaluation Committee (TEC) Request for Proposal, Measurement of Net-to-Gross (NTG) Factors for Ontario's Natural Gas Custom Commercial and Industrial Demand Side Management (DSM) Programs, Nov 1, 2013, page 10

process should not involve it making direct unilateral changes to the EC's scope of work, or to be the sole interpreter of Board Policy. Board Staff, in its role as the coordinator and overseer of the evaluation and audit process (as set out in the Board's August 21, 2015 letter, EB-2015-0245), is not the entity that should be in effect rendering a determination about how a Board Decision and Order should be interpreted and applied.

55. Board Staff certainly have the right to take a position before the Board in respect of issues and to adduce evidence in support of their position, but it is Enbridge's view that Board Staff's role does not contemplate a decision making function in respect of the interpretation of Board policies and rules. At a minimum, Board Staff should have required the EC to receive and reasonably consider the views, data and concerns of the utilities and to generate DSM program results with use of such information and to present the results using both interpretations. Presenting the results as two distinct deliverables would align with the original scopes of work, and separate the conflicting views and interpretations of Board policy from the study results themselves so the Board would have greater line of site on the issues to make a determination that is just, reasonable and consistent with the Framework and Guidelines.

56. In a memo provided by Enbridge to the EAC and Board Staff on June 14, 2017, Enbridge outlined its view that where a difference exists between any member of the EAC and/or Board Staff as to the interpretation and application of the Decision, such differences and the resulting impact on program results should both be presented in the evaluation results report which will ultimately then be filed with the Board. This would necessarily mean that where the EC has been directed to undertake an evaluation based upon an interpretation of a Board Decision which is in dispute, the EC should be required to undertake the evaluation using both interpretations so that there is a full record that is presented to the Board for

adjudication. Enbridge further suggested that it would then be open to each interested stakeholder to file evidence and make such argument each considers appropriate to support its interpretation of the Decision. Subsequent to receiving this memo, Board Staff did not and has not taken any action to respond to the concerns outlined. Throughout the 2015 DSM EM&V process, in an effort to seek clarity on OEB Staff's position on the application of NTG Study application to 2015 DSM program results, the utilities continued to raise concerns regarding the change to the NTG study scope of work and how the study outcomes would be applied (including at subsequent EAC meetings). During the October 2016 EAC meeting, OEB Staff committed to consider the matter and respond. However, as mentioned above, in an email sent to the two utilities on May 23, 2017, Board Staff formally communicated it had instructed DNV GL to retroactively apply the NTG Study ratios to 2015 DSM program results.

57. A further example of concern where the intended scope of work was not followed in the EC's execution of the NTG study is in regards to the determination and consideration of secondary attribution. This refers to the consideration of the longer-term effect of the program on participant decision making, which is particularly relevant to a mature program that has been in market for many years and where the utility has provided long term support of customers prior to current year projects. Resolution for consideration of secondary attribution in the NTG Study was documented by the TEC and DNV in the original scope of work such that it was agreed that while the primary objective of the free ridership estimation would be to capture the effect of the program(s) on the current project, the effect on the current project of prior and indirect program experience would be captured in a secondary, less rigorous question sequence. It was further communicated by DNV that the work plan would outline specifics for operationalizing this approach. Consideration of secondary attribution is also reflected in DNV's updated scope of

work for the CPSV/NTG verification.[20] In addition, the scoring methodology for secondary attribution outlined in the scope of work provides that the greater of this score and the primary attribution score would be used in determining the score for the participant. Also of note, though Enbridge provided comments in the development of the survey instrument, that multiple questions (in reference to "question sequence") above should be asked to capture this important component of utility influence on the customer, the EC did not incorporate this recommendation and limited the query to a single question. Subsequently, when the EC distributed the draft results of the free ridership evaluation, the EC asserted that while it had provided a quantified measurement of secondary attribution in its finding, it did not incorporate these values in the free ridership results. This is not in line with the original approach reviewed at the TEC and reflected in the original and updated scopes of work.

58. The omission of secondary attribution in the estimation of free-ridership values has a significant impact on findings. Leaving aside the issue of how and when NTG Study values should be applied, as well as other concerns Enbridge has with the findings, NTG ratios for Enbridge would be 10% higher when secondary attribution is correctly included in the value. Enbridge asserts that secondary attribution must be included in the NTG values in accordance with the original resolution with the TEC and DNV, and as outlined in both DNV's original and updated scopes of work.

[20] Measurement of NTG Factors and Custom Savings Verification For Ontario's Natural Gas Custom Commercial and Industrial DSM Scope of Work Ontario Energy Board, by DNV GL, dated December 14, 2016, page 44

Section 2 – Retroactive Application of NTG Study Results

The OEB Decision and Order

59. Enbridge is of the view that the Board's Decision and Order of January 20, 2016 as confirmed in its revised Decision and Order of February 24, 2016 (together the "Decision") in respect of the utilities 2015-2020 DSM Plans (EB-2015-0029/0049) along with the DSM Framework do not provide that NTG Study values are to be used in the determination of the Company's 2015 DSM program year results in a retroactive manner. Enbridge submits that the direction regarding retroactivity promoted by Board Staff is inconsistent with the Decision for the purposes of the evaluation of the Company's 2015 DSM results.

60. In the Board's Framework, the Board directed the gas utilities to "roll-forward their 2014 DSM plans, including all programs and parameters (i.e. budgets, targets, incentive structure) into 2015. … [Further]…the gas utilities should increase their budgets, targets and shareholder incentive amounts in the same manner as they have done throughout the current DSM Framework (i.e. 2013 updates to 2014 should now apply to 2014 updates to 2015)."[21] The Company complied and increased targets, budgets and the shareholder incentive in the exact same manner as the 2014 Plan. Similarly, these 2015 values involved a roll over and incorporation of the same inputs, assumptions and NTG values approved in the 2014 audit.

61. On January 20th, 2016, the Board released its Decision and Order which included the approval of Enbridge's 2015 budget, targets, metrics, scorecards and shareholder incentives as outlined in Enbridge's Multi-Year Plan. In its Decision, the OEB reconfirmed the direction provided in the Framework that "2015 would act

[21] EB-2014-0134, Report of the Board: Demand Side Management Framework for Natural Gas Distributors (2015-2020), December 22, 2014, p.37

as a transition year to the new multi-year DSM plans and that the gas utilities should carry forward and increase their 2014 DSM budgets in the same manner done from 2013 to 2014."[22] The Board stated that it "approves the gas utilities proposed 2015 DSM budgets. The OEB finds that the gas utilities have appropriately carried forward their 2014 DSM budgets into 2015."[23] The Board further stated in its Decision that it "approves Union and Enbridge's proposed 2015 metrics and targets for all scorecards. The OEB believes that it would be inappropriate at this time to make a change to the 2015 targets with the year completed."[24] Again, it is important to note that Enbridge's 2015 targets and scorecards were developed using the input assumptions and NTG factors that were used to determine 2014 DSM program results.

62. In the Company's view, since 2015 budgets and targets were the result of the Board's direction to roll over from 2014 budgets and targets, it is inappropriate to retroactively apply adjustments to a program year derived from a different set of input assumptions, including NTG values. Enbridge believes this is unreasonable and inconsistent with the spirit of the Board's instructions as part of the Multi-Year DSM Framework. In its 2015-2020 DSM Plan Decision, the Board confirmed this interpretation by stating that: "input assumptions and net-to-gross adjustment factors are finalized for a given year based on the previous year's final DSM audit."[25]

63. In section 5.2.6 of the Decision, the Board approved Enbridge's custom commercial and industrial offers as proposed. In addressing the custom commercial and industrial offers, the Board clearly stated that:

---

[22] EB-2015-0049, Decision and Order, January 20, 2016, p. 56
[23] EB-2015-0049, Decision and Order, January 20, 2016, p. 57
[24] EB-2015-0049, Decision and Order, January 20, 2016, p. 63
[25] EB-2015-0049, revised Decision and Order, February 24, 2016, p. 3

> The OEB does not expect the gas utilities to rely on a predetermined free ridership rate for the duration of the 2017 to 2020 term. In 2016, the free rider rates will be updated based on the results of the net-to-gross study and the annual evaluation process. Annually, the evaluation process will continue to inform the free rider rates for custom programs.[26]

64. The explicit reference made by the Board that it does not expect the utilities to rely on predetermined rates for the 2017 to 2020 term, aligns with the Board's introduction of the Target Adjustment Mechanism beginning in 2017 and deliberately leaves separate the 2015 and 2016 program years, for which the Board approved specified "fixed" targets. By deliberately not including 2015 or 2016 in its statement, the Board has provided a clear distinction with regard to treatment in these years. The Company submits that at the time of the Decision, the Board expected and anticipated that the NTG Study would be completed in 2016 in time to inform the development of programs in 2017 (in reality however the NTG Study is still incomplete). In other words, the NTG Study would be used prospectively. There is no indication in the Decision that the Board expected 2015 DSM results, which were based on a formulaic rollover, to be adjusted retroactively by an incomplete NTG Study released in October 2017.

65. Further support for the utilities' understanding on this issue can be found in the revised OEB Decision and Order dated February 24, 2016 in the Board's response to Union Gas' written comments submitted February 3, 2016. In these comments, Union Gas requested clarity on the treatment of input assumptions and NTG adjustments by explicitly requesting confirmation as follows:

> Union interpreted the OEB's Decision to mean that input assumptions and net-to-gross adjustment factors are finalized for a given year based on the previous year's final DSM audit…Given that the Board's Decision is effective for 2015 and based on the process outlined above, Union's 2015 results for the purpose of determining the 2015 DSM Incentive will be based on the same input assumptions and net-to-gross adjustment

[26] EB-2015-0049, Decision and Order, January 20, 2016, p.21

> factors used for setting Union's 2015 targets. These inputs were finalized in Union's 2014 DSM audit.[27]

In its revised Decision and Order, February 24, 2016, the Board provided the following confirmation:[28]

> *"The OEB confirms that Union's interpretation is correct."*

66. It is clear from Union's request for clarification that it was referring specifically to the manner in which the DSM incentive would be calculated. Like Union, Enbridge relied upon the Board's response and concluded that no adjustments to NTG factors as determined by the NTG Study would be applied in a retroactive fashion for the purposes of calculating the 2015 shareholder incentive.

67. Not only is the retroactive adjustment inappropriate based on the specific Direction provided by the Board, it stands to reason that such retroactive adjustments are inappropriate. At precisely the time the Board has tasked the utilities with doing as much as possible to mitigate carbon emissions, a clear and direct positive benefit derived from DSM activity, such retroactive adjustments change the "rules of the game" after the game has been played. Had the utilities known these input assumptions, and values could be changed to rearrange outcomes, the utilities would have been disincented to expend the degree of time and effort on Commercial and Industrial Custom projects as they did. Contrary to the Conservation Directive of the Government of Ontario, this would have resulted in higher past, current, and future, Cap and Trade offset purchase requirements for customers. Targets and results should be based on the same set of assumptions to ensure the Company can effectively plan, execute and deliver its business strategy. This business strategy is guided by the Direction and guidance provided

[27] EB-2015-0029/0049, Union Gas Written Comments RE: Decision and Order, February 3, 2016, page 2
[28] EB-2015-0049, Revised Decision and Order, February 24, 2016, p.3

by the Board in formulating its expectations for how the utilities set rates, and what activities they pursue. This is as true for DSM as it is for Compliance Planning.

68. Applying input assumption changes retroactively creates an unstable and unfair policy environment, which is contrary to the guiding principles enunciated in the Framework. Without question, the application of revised NTG values on a retroactive basis will materially disincent the utilities from attempting to achieve higher customer participation levels and minimizing lost opportunities (Guiding Principles 5 and 6). It will also result in the utilities being discouraged from pursuing commercial and industrial projects that often have long measure lives that produce long term energy savings contrary to Guiding Principle 8.

69. In addition, Enbridge submits that the retroactive application of NTG Study adjusted values is in direct conflict with Guiding Principle number 9 which provides that the amount of shareholder incentive will depend on a utility meeting or exceeding its DSM targets and will take into consideration the relative difficulty in achieving other goals. Enbridge was directed by the Board to carry over its 2014 DSM programs into 2015, which it did. Enbridge expended the effort and undertook the delivery of these programs successfully and relied upon the 2015 targets which were approved by the Board. Enbridge did everything in accordance with its approved 2015 DSM plan. Board Staff are now proposing that the shareholder incentive which was available in Enbridge's approved 2015 DSM plan is no longer available because a different set of values other than those which the utilities relied upon should be used to calculate the shareholder incentive. One could describe this as a bait and switch type of tactic.

70. A hypothetical example may prove helpful. Enbridge undertook its 2015 DSM programs using the inputs which were finalized in Enbridge's 2014 DSM audit as directed by the Board. If, for example, a commercial/industrial custom offer had a

free ridership rate of 25%, this value was rolled over into 2015 and Enbridge pursued commercial/industrial custom projects under the belief that this value would be used to calculate actual results. This is appropriate given that the targets which are intended to incent the utilities to aggressively pursue cost effective DSM were, in this case, based upon a 25% free ridership rate.

71. If for the purposes of determining the shareholder incentive, a 50% free ridership rate is used to calculate actual results, then it is wholly unfair to compare the results using a 50% free ridership rate with targets that were developed and relied upon using the 25% free ridership rate. Even if a 50% free ridership rate is the "correct rate" based upon a subsequently completed study, what is being proposed by Board Staff is that this "correct rate" should be used in one half of the equation and that the wrong free ridership rate (i.e. 25%) should be used in the other half of the equation for the setting of targets. The methodologies are different and thus you have an apples and oranges comparison which Enbridge submits is wholly inappropriate. If actual results values are to be revised retroactively, then the values used to develop targets in the first instance should be similarly adjusted so that there is an apples to apples comparison.

Section 3 – Concerns Regarding the NTG Study

Best Practices in NTG Estimation

72. Enbridge is of the view that the EC's NTG Study did not incorporate best practice approaches. Consequently, Enbridge and Union Gas recently consulted Navigant Consulting, Inc. and Apex Analytics, LLC (the "Navigant team") to undertake a jurisdictional review to investigate NTG policies and practices, examine recent attribution policy developments, and explore best practices utilized in the consideration, assessment and application of NTG values. Their report is submitted in this application at Exhibit B, Tab 6, Schedule 1.

73. The report provides insights regarding the current NTG landscape; information that should inform Ontario stakeholders in relation to NTG in evaluation, program planning, use in measuring progress toward savings targets, and in determining shareholder incentives. Included in the Navigant team's findings are case studies of 3 leading jurisdictions: California, Massachusetts and Illinois. These jurisdictions were selected because, similar to Ontario, they have a long history of large-scale utility efficiency programs and have addressed many of the same issues regarding NTG brought to the forefront in the 2015 EM&V process. Upon review of the Navigant report, it is clear that the NTG study and more broadly the evaluation structure in place to facilitate such a study did not reflect best practice standards in place in leading jurisdictions.

74. The case study review produced some repeated overarching themes and cross-cutting findings highlighting best practice approaches to NTG estimation and methods: [29]

- In all three cases, structures have been developed that reduce the influence of after-the-fact (i.e., retrospective) application of NTG estimates (neither Massachusetts nor Illinois apply retrospective NTG estimates for determining shareholder incentives and California's new structure has significantly reduced the effect of retrospective application of NTG estimates by utilizing retrospective NTG only for select programs and by making it only one of four factors that are used in incentive determination).

- Where the purpose for using net savings has been established as a means of aligning utility goals with ratepayer value, in this context, NTG analysis provides information to inform energy efficiency investment and program planning. At the same time experts in these states reported that using net savings puts pressure

[29] Net-to-Gross Policies: Cross-Cutting Jurisdictional Review by Navigant Consulting, Inc., and Apex Analytics, December 14, 2017

on the accuracy of NTG evaluation efforts, especially if NTG were to be applied retrospectively to assess savings results and determine shareholder incentives. Instead, applying NTG estimates prospectively reduces uncertainty by eliminating the risk of applying retroactively, a different NTG ratio than was assumed in program planning.

- The process undertaken to finalize NTG estimates to be used in incentive calculations, establishing targets and in informing program design involved much more than simply accepting the results of a study. In all states, stakeholders worked together to review, challenge and modify initial estimates from EM&V studies, for example aiming to arrive at a consensus value considering relevant issues raised and factors to be considered. All three states had an established collaborative, transparent stakeholder process which aims to seek agreement among stakeholders as part of the finalization of NTG estimates, particularly in the case of self-report survey methods.

- In addition to the objective of transparency and review in the determination of final NTG estimates which serves to improve confidence for all stakeholders in the NTG estimation process, all states have adopted established agreed upon approaches incorporating pre-defined methods including agreement on self-report survey instruments/questions and scoring algorithms (which incorporate multiple influence factors – program, trade ally and market based), tested through sensitivity analysis. Experts reported the prospective application of results, combined with the consistency of the pre-defined methods and a transparent collaborative stakeholder process has created more certainty and confidence for stakeholders regarding the actions needed to meet targets, as well as allowed for an increased focus on the continuous improvement of programs.

- All experts noted concerns with self-report methods but said that the primary method for custom project NTG is self-report survey methods due to the unique nature of commercial and industrial (C&I) custom projects. However, experts noted the following best practices approaches are used to improve accuracy and confidence:
  - Fast Feedback – involves conducting the survey as soon as possible after a project is completed, where respondents are asked about influencing factors to program participation near the time of participation (e.g. within 3 months of completion). This approach helps mitigate recall bias and increases the likelihood of evaluators contacting an informed person who was involved with the project is question.
  - Sensitivity Analysis with full transparency regarding participant scoring has been used in all states, particularly when the survey batteries are first developed and tested. This is particularly important since different, but still reasonable assumptions in translating question responses in NTG scores can result in very different NTG value determinations.
  - Triangulation: The perspective of vendors/business partners is collected in all states on a project-by-project basis. Experts noted the challenge for participants in differentiating the attribution of any individual influence on decision making. Triangulation which includes surveying vendors/trade ally who are often a key to program delivery working with the utility, is used in best practice in the NTG estimation including to inform the relative influence of multiple program influences.
  - Other best practices noted included: incorporate multiple factors (program influence as well as non-program influence) in NTG scoring and ensuring questions are fully vetted and gaining insight into the project story from the

participant and meeting with implementation personnel familiar with the project.

- Experts in all states agreed with the approach adopted in Massachusetts which incorporated the inclusion of previous program influence (i.e. influence that builds over time when a program cover multiple years) in the scoring algorithm. Further, experts in all cases noted there are improvements that could be made in estimating spillover and market effects.

Selected NTG Study Estimation Issues

75. Over the course of the 2015 EM&V process, Enbridge provided extensive verbal feedback and written comment to the EC, the EAC and Board Staff. In some instances feedback was offered to highlight opportunities for improvement, to clarify details for the EC or to seek clarification on particular items, as well as to raise concerns Enbridge had with the scope or approach of 2015 EM&V activities. In the case of the NTG study, from the scoping phase through the planning and execution of activities many of the comments that Enbridge put forth were not addressed, and in some cases where inconsistencies were identified by Enbridge, such issues were often not appropriately resolved.

76. Enbridge is of the view that the survey instrument employed by the EC focused the customer largely on the program's provision of customer incentive payments and did not sufficiently probe for the customer's impression of all the services, support and value provided by the utility. "If a survey is conducted 1 year or more after participation in a program, the respondent may not recall all the features of the program and all the assistance provided. Instead, respondents may focus narrowly on the influence of the rebate or incentive payment."[30] Utility support and therefore

[30] Memorandum: Discussion of Selected NTG Estimation Issues by Navigant Consulting, Inc., and Apex Analytics, December 14, 2017, page 14

influence can be part of any number of customer engagement activities, for example, site assessment, facility audits, project feasibility studies, marketing communications, case studies, workshops and education events and generally through on-going customer relationship development and support over many years. Limiting assessment of these varied influence factors puts the determination of the NTG scoring in question.

77. As discussed previously, concerns regarding the EC's decision to not factor secondary attribution was raised repeatedly by Enbridge but dismissed. Enbridge also provided considerable feedback on the survey instrument, much of which was similarly not incorporated.

78. Importantly, the NTG Study scope of work specified that the EC was to interview Enbridge's program advisors in order to ensure clear understanding of the program and the role of Enbridge consultants in working with the customers. In fact, DNV had outlined early in the process that this was an important step to appropriately frame questions in the survey process. As detailed on page 33 of the scope of work: "Program energy advisor interviews will be scheduled after submission of the draft SOW. These interviews will focus on the specifics of program interactions with customers. The intent of the interviews is to ensure that the FR framing in the IDIs [in-depth interviews] and CATI [computer assisted telephone interviews] covers the range of program activities that may have influenced decisions to implement projects." Enbridge highlighted that this activity was not yet completed when the EC was preparing to initiate surveys with customers but was told a decision was made (by the EC and/or OEB Staff), to eliminate these interviews. Enbridge communicated its concern with this omission to no avail. The Company presumes because the NTG effort was well delayed at this point, there were pressures to expedite the process and therefore this activity was omitted.

79. Enbridge continues to have questions regarding the scoring approach employed by the EC in the NTG Study. The scoring process involves the translation of survey responses into NTG scores or values and includes the application of a scoring algorithm and the introduction of parameters to apply judgement in the survey. In addition to delivering their *Net-to-Gross Policies: Cross-Cutting Jurisdictional Review* report summarized above, the Navigant team also provided the utilities with a companion *Memorandum: Discussion of Selected NTG Estimation Issues* ("Memorandum"). This document is included in this submission at Exhibit B, Tab 6, Schedule 2. The Memorandum clearly identifies the importance of undertaking a sensitivity analyses to test the scoring methodology to better understand the implication that assumptions used in translating survey responses into NTG scores can have on NTG values:[31]

> The scoring algorithm is central to any resulting NTG estimates. As a result, it is important that the algorithms be as transparent as possible and undergo a stakeholder review process to build confidence in the approach. A process that allows for discussion of the scoring algorithms, includes sensitivity analyses to assess robustness, and is as transparent as possible is important for producing NTG values that will have buy-in from stakeholders.

80. Enbridge, nor the EAC, was involved in any sensitivity analysis undertaken in the EM&V process. There was no such activity outlined in the scope of work and the Company is unaware of any such analysis conducted on the NTG Study. This contributes to Enbridge's view that the NTG study did not follow best practice and further reduces the confidence the Company has in the results.

81. Among a number of NTG estimation issues about which the utilities requested the Navigant team's perspective, the Memorandum provides some comparative examples to illustrate how slightly different assumptions made in a scoring algorithm can have significantly different outcomes in the determination of NTG

---

[31] Memorandum: Discussion of Selected NTG Estimation Issues by Navigant Consulting, Inc., and Apex Analytics, December 14, 2017, page 3

values. For example, the Memorandum provides a comparison of how the calculation of a NTG score for the same response to a specific question has been scored very differently across Massachusetts and Ontario. In a few specific cases presented in the Memorandum, the question relating to a timing probe is the same in both Massachusetts and Ontario. The analysis refers to the EC's report which provides Example Attribution Calculations in Table 8-86 [32] of selected examples of scores translated from survey responses. In one example from the NTG study, where a respondent indicates that they would have undertaken the project "24 months later" if the utility program had not been offered, this results in an NTG of 31%. For the same question, and the same "24 months later" response, the resulting NTG score in Massachusetts is 50%. This is a 60% increase in the NTG score (i.e., 50% NTG/31% NTG) on a single question due to a different scoring algorithm. A second example, which compares the same response provided across each jurisdiction to an efficiency question, results in a NTG score of 38% applying the EC's algorithm in Ontario, however in Massachusetts, the NTG score would be 50%. The exercise serves to illustrate how differences in the assumptions/parameters utilized in the scoring algorithms can influence NTG values, even where the questions are very similar.

82. Acknowledging Navigant's comparison of Ontario to Massachusetts in the above example, a further review of the EC's determination of a proxy spillover value is fitting. As outlined previously, late in the 2015 EM&V process, prior to the September 27, 2017 EAC call, the EC had been asked to conduct secondary source research to identify an estimation of spillover that might reasonably be applied to the utilities' 2015 DSM programs as an estimate and then proceed to finalize the 2015 program results verification.

[32] EB-2015-0245, Ontario Gas DSM Evaluation Contractor 2015 Natural Gas Demand Side Management Custom Savings Verification and Free-ridership Evaluation, October 12, 2017, page K-1

83. The EC recommended a finding of 3.4% spillover from a study conducted in Massachusetts by Tetra Tech.[33] The EC proposed this was the most applicable value for the Ontario DSM programs because: [34]

- Massachusetts has a similar climate to Ontario's major population centers, so it is likely that similar measures are being implemented
- The spillover value is specifically for custom gas C&I measures, which is the same program type
- The programs in Massachusetts and Ontario are mature and in leading jurisdictions
- The Massachusetts study looked at both "like" and "unlike" spillover
- The rate is within the anticipated range of results expected for spillover from custom gas C&I programs, not an extreme outlier
- The study is relatively recent, from 2014-15.

The EC further outlined some differences from the Ontario spillover study as follows: [35]

- It only *quantifies* (provides a savings estimate for) like spillover, not unlike spillover.
- The study was conducted on customers who had participated in the program 15-27 months prior, not four or five years ago. This provides for less time since the program measure for spillover to occur.

---

[33] Tetra Tech (Revised August 10, 2015). "2014-15 Commercial and Industrial Natural Gas Programs Free-ridership and Spillover Study" for National Grid, Eversource, Unitil, Berkshire Gas, Columbia Gas of MA, and Liberty Utilities.
[34] Ontario Gas DSM Evaluation Contractor 2015 Natural Gas Demand Side Management Annual Verification, Ontario Energy Board, October 12, 2017, Appendix N, page N-1
[35] Ibid., Appendix N, page N-1.

84. What is particularly interesting to Enbridge upon review of the Tetra Tech study is that in addition to determining a 3.4% spillover value (that the EC has deemed is reasonably appropriate and applicable to the Ontario DSM programs), the Tetra Tech corresponding free-ridership determination provided in the Massachusetts study was 15.7%. With similar consideration for the reasons outlined above by the EC for why the deemed spillover value of 3.4% is an appropriate proxy for application in Ontario, Enbridge concludes it would be equally reasonable to determine that the free-ridership rate should be likewise applicable. However in the EC's NTG study it is suggesting a free-ridership rate of 73% for Enbridge. In addition to the many other factors highlighted in this evidence that put the reasonability and credibility of the EC's NTG study into question, the starkly disparate values for Commercial/Industrial custom free-ridership determined in the Massachusetts and Ontario studies further erodes Enbridge's confidence in the EC's NTG study.

85. The Evaluation Contractor highlights that the Massachusetts spillover estimate falls within the anticipated range, and is not outlier. This point cannot be made however for the free-ridership estimate proposed by the EC in the NTG study. In 2013, Navigant Consulting was contracted by the TEC to conduct a jurisdictional review of free-ridership and spillover values. A key finding from the third-party review of 42 jurisdictions revealed that "while the dispersion of net-of-free ridership values is quite large, ranging from 21% to 100%, the majority of values appear to 'cluster' between 40% and 90%."[36] Only two of the values detailed in the review were below 40%. However, a proposed free-ridership value of 27% as put forth in the EC's NTG study result falls significantly outside an already wide clustering of NTG values indicated in the 2013 Navigant jurisdictional review. Enbridge is of the view

[36] Custom Free Ridership and Participant Spillover Jurisdictional Review Prepared for: Sub-Committee of the Ontario Technical Evaluation Committee, May 9, 2013, page 18.

this observation provides further evidence that the NTG study findings should be questioned.

86. The December 14, 2017 Navigant Memorandum further explores a concern Enbridge had shared with the EC relating to questions in the survey design which aim to assess partial free ridership by probing the customer about the efficiency they might have undertaken in the absence of the program. For example a question asks "*Without <the program>, would you have installed <measure> that was "standard efficiency on the market at that time," or "between standard efficiency and the efficiency that you installed?"* Only if the respondent knows or understands what standard efficiency is to compare to the higher efficiency equipment/features installed through the program are they able to provide an informed answer to these types of questions. The application of baselines to gross savings calculations and the consideration of baselines in NTG calculations is complex, particularly in situations of early replacement. The Memorandum again points out "the only way to really address concerns about potential biases in the response to NTG questions that have a baseline assumed (e.g., installing equipment above standard practice) is to perform sensitivity analyses."[37] The Memorandum further points out that recent research has shown responses to these types of partial free rider questions can be highly variable, and therefore again highlights the need for further analysis.

87. In regards to the determination of NTG estimation for the Run-it-Right ("RiR") offer, Enbridge does not believe an appropriate approach was undertaken. The EC employed much the same survey instrument utilized for the purposes of the Commercial/Industrial custom NTG participant queries. RiR is uniquely different from the Company's custom offer and one that requires a multiple year

[37] Memorandum: Discussion of Selected NTG Estimation Issues by Navigant Consulting, Inc., and Apex Analytics, December 14, 2017, page 18

participation on the part of the customer. Enbridge is of the view that determination of net savings for this type of program necessitated an appropriate, tailored and applicable evaluation.

88. The participant in Run-it-Right is enrolling in a process not undertaking a project. Upon agreement and enrollment, participation starts with an investigation agent who visits the facility to undertake a site assessment/audit to identify, from a list of eligible measures, recommended low cost recommissioning/ operational improvement activities that should be undertaken by participants with the goal of improving the operation of their buildings in terms of energy usage. The customer must agree and proceed to implement recommended improvements and allow Enbridge to monitor gas usage for 12 months and the offer provides access to an Energy Management Information System ("EMIS"). Regression analysis is completed 12 months hence and savings are calculated net of weather normalization. Framing "projects" to participants in RiR does not make sense. Customers were recruited for enrollment in a process through the RiR program with the objective that Enbridge would recommend the improvements to undertake to reduce consumption, i.e. directly influence the customer to take action. The consideration of these types of unique programs is outlined in the Memorandum[38]:

> …there are a growing number of programs that are trying to move away from paying out large incentives. Instead, there is greater emphasis on engagement, information, and business case development—all of which support a more favorable environment for investments in EE. For these programs, it is often important that the survey introduces the ways support was provided through the program. This would include making sure that program training, analysis, and support are described to the participant. These can be particularly difficult for the respondent to recall if the survey takes place 1 year or more after participation. A program driven by financial incentives to induce participation has one major event (i.e., the payment of incentives) that the respondent can easily recall. More sophisticated programs that work to engage and support customers in making EE investments can require different survey designs to capture these non-incentive influence factors.

[38] Memorandum: Discussion of Selected NTG Estimation Issues by Navigant Consulting, Inc., and Apex Analytics, December 14, 2017, page 18

89. A further concern shared by Enbridge regarding the EC's NTG findings related to the requirement, as outlined for the NTG study, that the sampling methodology should be designed to achieve a 90/10 precision target ("90/10" precision is a statistical standard for which there is 90% confidence that sample results are within +/- 10% relative precision). However the relative precision of some of the ratios did not come close to meeting this expectation. Enbridge further pointed out concerns with the error ratios identified and that such values did not instill confidence in the results. The EC subsequently acknowledges this concern in that they include the following recommendation in the findings outlined in the final Annual Verification Report. "Error ratios from the results provided in this report should be used to inform sample design for future evaluation years… [Further]... Better defined error ratios for the measures in the programs will allow more efficient sample." [39]

90. The preceding examples illustrate Enbridge's concerns with the NTG study and contribute to Enbridge's view that the NTG Study does not follow best practices. The concerns articulated regarding the EC's approach to this evaluation serve to illustrate that the Company does not have confidence in the EC's determinations.

Broadly Identified Issues with Self Report NTG Assessment Approaches

91. Enbridge's concerns with the NTG Study in respect of the appropriateness and reasonableness of their retrospective application to 2015 program results is further compounded given consideration for the well documented failings of self-report NTG approaches.

92. Many of Enbridge's concerns with the credibility of the NTG Study results are reflected in the broader industry discussion on the risks and merits of the self-reporting method for assessing NTG values. The self-report method, by its nature,

[39] Ontario Gas DSM Evaluation Contractor 2015 Natural Gas Demand Side Management Annual Verification, Ontario Energy Board, October 12, 2017, page 81

presents a host of methodological risks and failings that put its results further into question.

93. In its September 1, 2017 submission, as part of the 2015-2020 DSM Mid Term proceeding the Company engaged Dr. Jane Peters, of Research into Action, Inc. (“RIA”), to review and analyze current literature relating assessment methods for NTG. RIA’s report, *Review and Analysis of Net-to-Gross Assessment Issues for Natural Gas DSM Custom C&I Programs*, is included in this submission at Exhibit B, Tab 6, Schedule 3

94. RIA acknowledges self-report surveys as being very commonly employed in the industry, in part due to their “low cost and ease of administration”. However, the report primarily cautions that the self-reporting method inherently suffers from several sizeable challenges that put results into question. These concerns included but are not limited to the following:

- Inaccuracy in attempts to have respondents attribute influence to various sources, including the impact of respondents’ own bias to provide socially desirable answers which reflect higher levels of environmental conscientiousness than may in fact be the case;

- Difficulty in reporting the counterfactual, hypothetical alternative which is where respondents are asked to speculate on the hypothetical scenario of what they might have done if the program in question hadn’t existed. Such conjecture is influenced by many factors including what is relevant at the time of the surveying when it is expected that energy efficiency is top of mind;

- Respondents’ tendency to rationalize past decisions in a way that avoids contradiction between their actions and their stated attitude regarding energy efficiency in addition to the tendency for respondents to lean to providing socially desirable replies;

- Evaluation approaches regularly fail to tease out all of the direct and indirect pathways through which programs influence customer behaviour. By way of example, Enbridge's efforts working with contractors and installers may influence many end-user decisions without those customers' direct knowledge of such influence taking place;

- Respondents' difficulty remembering the specific intentions, motivations or other influences which underpinned their past energy efficiency decision. Where numerous public policies and market interventions influencing energy efficiency exist simultaneously, it is likely impossible to extract the influence of a single program. This difficulty increases as time elapses between the decision point and evaluation efforts, with increased difficulty remembering subsequently increasing the likelihood that customers defer to existing biases to internalize energy efficiency decisions as self-originated.

95. Enbridge submits that the distorting effects of the above factors on survey results are increasingly exacerbated in Ontario as a growing number of independent market entities promote energy efficiency activities, engaging customers and dispensing financial incentives. The Company is of the view that the expectation that the self-reporting method will accurately tease out the influence of a single program has become increasingly doubtful.

96. In its summary recommendations, RIA highlights the importance of accurately assessing and including spillover estimates as well as market effects assessments, where feasible. Having delivered DSM to customers since 1995, Enbridge believes any such NTG assessment should indeed include a comprehensive estimation of the immediate and longer-term cumulative effects of the Company's lengthy

relationship with its customers, educating them and advocating for energy efficiency for over 2 decades. As explained in this submission, the self-report method selected by the EC remains incomplete having not concluded the spillover component. In addition the determination failed to incorporate a measure of quantification of the important longer term influence which should have been explored and factored through the secondary attribution consideration that was omitted by the EC.

97. RIA also highlighted the importance of undertaking any assessment of free-ridership as close as possible to a project's implementation. As documented, despite having an endorsed work plan in March 2016 with the TEC, as a result of the evolution of the 2015 EM&V process, the EC did not interview respondent until late January – April, 2017, in some cases more than 2 years after projects were concluded. Enbridge reiterates that this considerable delay only further weakens the reliability of respondents' comments and, in the end, the study's results.

98. Further, RIA recommends the use of multiple methods to triangulate NTG estimates. As determined in the EC's methodology, Enbridge is of the view that business partners for example (contractors; third party vendors) were not adequately engaged by the Evaluation Contractor to corroborate or clarify customer opinions of Enbridge's influence on their decision. Enbridge's long standing practice working with contractors and installers to help influence end-user decisions undoubtedly occurs at times without customers' direct knowledge of such influence taking place.

Section 4 – Basis for Application

99. The foregoing evidence identified issues with the EM&V process and the accuracy of the NTG Study results in a number of material areas. These concerns alone support the determination that for the Clearance of 2015 DSM Deferral and Variance Accounts, the retroactive application of any revised NTG values for 2015 is inappropriate and unfair. It is also contrary to the Board's findings, and hence should not to be applied to 2015 results. Enbridge is therefore applying for approval for the Clearance of its DSM Deferral and Variance accounts based on the application of the DNV CPSV results (as well as the verified results determined for all other evaluated program results) with the application of the same 2015 NTG values that were inherent in the formulation of 2015 targets (each rolled over from 2014, in accord with the Board's instructions) consistent with 2014 audited results.

100. Should the Board decide that the EC's NTG Study results should be applied to the Company's 2015 DSM results, then Enbridge submits that the Board should similarly determine that a corresponding adjustment should be made to the 2015 targets. The Company however does not believe that this was the intention of the Board given its clear direction in respect of the establishment of budgets and targets for 2015, namely that these would be established by a roll over from 2014. Similarly, Enbridge is of the view that the Board's Direction on the application of NTG Study values in a retroactive manner is also clear: it should not occur.

101. What the Company finds particularly troubling about DNV's findings and results, are, the implications that they have at a much broader level. If the results are to be believed, then the Board should be satisfied that no further Compliance Planning efforts to mitigate carbon and promote energy

conservation are required given that three quarters of the utilities commercial industrial customers are apparently undertaking DSM activities on their own without need of any input from the natural gas utilities With the Government's announced commitment to emissions targets, and the importance of energy efficiency activity in achieving these goals, this is an incongruous conclusion to draw. The real world is of course not so black and white. There can be no question that the effort of the natural gas utilities supporting conservation through promotion of custom conservation solutions advances the government's goal of conservation first and carbon emissions reductions. In the real world, a commercial or industrial customer may be able to look to a number of sources for advice, engineering design and funding. The fact that all of these sources exist highlights the importance of energy conservation and carbon emission reductions to the Government of Ontario. It does not, as the NTG Study suspect free-ridership values falsely suggest, indicate a declining need for such DSM programs.

102. As well, applying the NTG Study values to future DSM Plan targets will result in significantly reduced targets all of which may cause concern with both rate payers and the Government of Ontario who all seek a material decrease in carbon emissions in the short term. As outlined in its submissions for the Mid-Term Review, as the level and pace of activity continues to ramp up as the Province orients itself to meet its emissions targets by spending Cap & Trade Funds, then the attribution of utility activity can only wane, resulting in even higher Free Ridership rates. Compliance Planning, and the mitigation of carbon related expenses, are predicated on gross volumes. In other words, the inevitable outcome would be less utility activity and higher carbon related Cap & Trade expenses, both of which will result in higher rates for ratepayers.

<u>Relief Sought through this Clearance Application</u>

103. For the reasons set out in this Application, Enbridge respectfully requests that the Board make the following findings, determinations and orders:

a) Approve the CPSV portion of the EC report impacting custom commercial and industrial results, and similarly approve all other evaluated results from the balance of the 2015 program;

b) Reject Board Staff's proposal to retroactively apply NTG Study values developed in 2017 to the Company's 2015 commercial and industrial custom energy savings claims (including Run-It-Right);

c) Approve Enbridge's deferral and variance accounts balances for DSMVA, LRAMVA and DSMIDA values as outlined in Exhibit A, Tab 1, Schedule 3, Table 1 and restated below;

d) Direct Board Staff to work with the EAC to finalize the NTG Study by undertaking best practice approaches by requiring the EC to: (i) update the NTG study findings to include secondary attribution and spillover results; (ii) undertake a sensitivity analysis on the results for further review; and, (iii) act as a facilitator in respect of the determination of an appropriate NTG value through a collaborative, transparent, negotiated stakeholdering process in line with best practices in other leading jurisdictions; and,

e) Address the concerns outlined by the utilities regarding the 2015 verification process in respect of objectivity, transparency and collaboration by accepting and approving for future use the draft Charter filed by Union which outlines the roles and responsibilities of members of the EAC and includes a process which promotes consensus-based collaboration and decision-making.

104. The resulting impact to each of the LRAM, DSMVA, DSMI are shown in the table below:

2015 DSM Deferral and Variance Accounts and Balances	
Demand Side Management Variance Account (DSMVA)	$ 825,460
Demand Side Management Incentive Deferral Account (DSMIDA)	$ 10,077,695
Lost Revenue Adjustment Mechanism Variance Account (LRAMVA) (Reimbursable to Ratepayers)	$ (71,829)
Total Amount Recoverable	$ 10,831,326

2015 Demand Side Management Annual Report

December 18th, 2017

ENBRIDGE

Preface

In preparation of Enbridge's Application for the Clearance of 2015 DSM Deferral and Variance Accounts, EB-2017-0324, the 2015 Demand Side Management Annual Report has been updated from its original draft, following the release of the Evaluation Contractor's (DNV-GL) final 2015 Natural Gas Demand Side Management Annual Verification.

In accordance with details provided in Enbridge's Application for Clearance of 2015 DSM Deferral and Variance Accounts, Enbridge's 2015 Demand Side Management Annual Report reflects all 2015 verified program results as presented in the Evaluation Contractor's Annual Verification report with the exception of the Net-to-Gross ("NTG") Study findings.

Table of Contents

List of Tables

Executive Summary

Enbridge Gas Distribution (Enbridge, or the Company) summarized its 2015 DSM Plan in the 2015-2020 Multi-Year DSM Plan (EB-2015-0049), filed on April 1st, 2015. The Company's 2015 DSM plan was outlined consistent with the transitional provisions set out by the Board in the Report of the Board: Demand Side Management Framework for Natural Gas Distributors (2015-2020), published December 22nd, 2014 (EB-2014-0134).

In its Decision and Order, published January 20th, 2016, the Board agreed and determined that Enbridge reasonably interpreted the DSM Framework and subsequently approved the Company's 2015 budget, metrics and targets for all scorecards, shareholder incentive amounts, and incremental budget as filed.

The Company is pleased to report that in the 2015 DSM program year, the portfolio generated total net annual natural gas savings of 49.0 million cubic meters (m^3) or 826.2 million net lifetime (cumulative) cubic meters (CCM). These savings are a direct result of the Company's ongoing efforts delivering the Resource Acquisition and Low Income programs. Natural gas savings attributable to Market Transformation program delivery are not captured in these totals since results for this program are not measured on the basis of cubic meters (m^3) or lifetime (cumulative) cubic meters (CCM) saved.

As outlined in the Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2020), submitted by the Board on December 22nd, 2014 (EB-2014-0134), the Board calls for application of a Total Resource Cost (the TRC-Plus) test as well as the introduction of the Program Administrator Cost (PAC) test to screen for cost-effectiveness of programs. In 2015, the portfolio demonstrated cost-effective program delivery based on positive results from both the TRC-Plus and PAC screening tests. The TRC-Plus ratio for the Resource Acquisition program was 3.12, while the TRC-Plus ratio for the Low Income program was 1.88 – both well above cost-effectiveness screening thresholds. Overall the 2015 portfolio had a TRC-Plus ratio of 2.95. The PAC ratio for the Resource Acquisition program was 5.21, while the PAC ratio for the Low Income program was 2.00. The 2015 portfolio had an overall PAC ratio of 4.47.

The Company continues to be proud of its accomplishments in DSM and is pleased it was able to demonstrate successful results relative to 2015 Board approved targets across the range of the various offers.

Table ES.0 2015 DSM Portfolio Results

Program	Annual Net Gas Savings (m3)	Cumulative Net Gas Savings (m3)	Budget	2015 Spending	TRC-Plus Ratio	PAC Ratio
Resource Acquistion			$14,443,790			
Residential	6,762,791	102,415,214		$9,362,295	2.24	1.94
Commercial	25,646,715	450,722,741		$6,221,724	3.39	10.78
Industrial	12,289,466	180,990,879		$2,166,706	6.15	15.45
Overheads			$4,731,485	$5,639,080		
Total Resource Acquisition	**44,698,972**	**734,128,834**	**$19,175,275**	**$23,389,805**	**3.12**	**5.21**
Low Income			$6,864,090			
Part 9 (Single Family)	1,129,070	28,067,264		$4,444,616	1.06	0.96
Part 3 (Multi Family)	3,143,515	63,969,353		$2,111,746	3.20	4.76
Overheads			$517,988	$617,349		
Total Low Income	**4,272,585**	**92,036,617**	**$7,382,078**	**$7,173,710**	**1.88**	**2.00**
Market Transformation			$4,890,900			
SBD Residential	n/a	n/a		$2,032,022	n/a	n/a
SBD Commercial	n/a	n/a		$890,464	n/a	n/a
Home Labelling	n/a	n/a		$121,241	n/a	n/a
Overheads			$1,353,687	$1,613,352		
Total Market Transformation	**n/a**	**n/a**	**$6,244,587**	**$4,657,079**	**n/a**	**n/a**
Grand Total *	**48,971,556**	**826,165,451**	**$32,801,939**	**$35,220,594**	**2.95**	**4.47**

* Budget and Spendings amounts do not include the $4.92 million incremental budget nor the spending in 2015 against that budget. Incremental budget/spending is detailed in section 10.3 of this report

Overall the Resource Acquisition program contributed 734.1 million net CCM in natural gas savings. Resource Acquisition offers targeted to the Commercial and Industrial sectors achieved net gas savings of 450.7 million and 181.0 million CCM respectively. The Residential home retrofit offer which has seen excellent growth since its launch in mid-2012 contributed 102.4 million net CCM savings and reached 5,646 households.

The Low Income program delivered 92.0 million net CCM gas savings in 2015. Results for both the Single Family (Part 9) offers, which provided 28.1 million net CCM and for the Multi-Residential (Part 3) offers which contributed 64.0 million net CCM exceeded targets set out in the 2015 scorecard for the Low Income program.

Market Transformation offers continued to demonstrate substantial results in 2015, reaching or exceeding weighted scorecard upper targets for all three of the Savings by Design Residential, Savings by Design Commercial and Home Labelling offers.

DSM results for 2015 were achieved with total spending of $35,220,594 million. The OEB approved budget for 2015 as per the Board's January 20th, 2016 Decision was $32,801,939. In addition, the Board approved an incremental budget of $4,920,291 as outlined in the Company's 2015 DSM Plan. Incremental spending totalled $559,378 in 2015; this spending is detailed in Section 10.3 of this report.

The maximum DSM shareholder incentive available for the 2015 program year is $11,089,624. The determination of the Company's incentive is based on 2015 DSM performance in relation to the weighted scoring approach. The resulting DSM Shareholder Incentive earned by the Company for 2015 is $10,077,695.

The Lost Revenue Adjustment Mechanism Variance Account (LRAMVA) is utilized to true-up the lost distribution revenues associated with DSM activity relative to what was included in the forecast for rate-setting purposes. The LRAM amount calculated for 2015 is $71,829 to be refunded to ratepayers.

The DSM Variance Account (DSMVA) is utilized to track the difference between DSM spending in 2015 and the amount already built into rates. This amount totalled $825,460 to be recovered from ratepayers.

Table ES.1 2015 DSM Results Summary

2015 DSM Results Summary	
Net CCM Savings	826,165,451 m^3
DSMIDA amount recoverable from Ratepayers	$10,077,695
LRAMVA amount to be paid back to Ratepayers	($71,829)
DSMVA amount recoverable from Ratepayers	$825,460

* The DSMVA represents the difference between the 2015 budget already built into rates and the 2015 spending including incremental spending.

1. Introduction

Following a directive from the Ontario Energy Board, (EBO 169-III) in 1995, Enbridge launched a suite of Demand Side Management (DSM) programs and activities to help its customers reduce their demand for natural gas. Demand Side Management is defined as "…actions taken by the utility or other agencies which are expected to influence the amount or timing of a customer's energy consumption."[1]

Enbridge's DSM programs are developed with stakeholder consultation and are funded through Board approved Enbridge Gas Distribution rates. In 1999, Enbridge was granted Board approval to receive a financial incentive for DSM activities by way of the Shared Savings Mechanism (SSM), which was replaced by the Demand Side Management Shareholder Incentive in 2011.

The continuing need for DSM efforts in the province of Ontario was outlined by the Board in the Demand Side Management Guidelines for Natural Gas Utilities (the Guidelines), published June 30th, 2011, and again in the Report of the Board: Demand Side Management Framework for Natural Gas Distributors (2015-2020), published December 22nd, 2014 (the Framework).

The Framework sets out three primary goals to guide the utilities' DSM portfolios:

- Assist consumers in managing their energy bills
- Promote energy efficiency and create a culture of conservation
- Avoid costs related to future natural gas infrastructure investment

The Framework also provides direction for DSM programs and outlines the proposed weighted scorecard approach to measuring DSM performance.

The Company has had significant achievement in results since Demand Side Management was introduced to its customers. From 1995 to 2014,[2] Enbridge's DSM programs have collectively reduced customer consumption by 9.6 billion cubic metres of natural gas, which is roughly enough natural gas savings to serve nearly four million homes[3] for one full year. In emissions, this translates to a reduction of 18 million

---

1 EBO 169 Appendix B, Glossary of terms, pg. 4
2 Subject to 2014 Clearance of Accounts proceeding (EB-2015-0267) before the Ontario Energy Board
3 Assumes a residential customer using 2,400 m3 per year to heat their home and water

tonnes[4] of carbon dioxide emissions, which is the equivalent of removing 3.5 million[5] cars from the road for one year.

Enbridge is pleased to continue to offer DSM programming through the Board approved 2015-2020 Multi-Year DSM Plan to help its customers reduce their energy bills, and to provide support for the Province's greenhouse gas reductions emissions targets.

The 2015 Annual Report on Enbridge's Demand Side Management programs provides a summary of the results achieved over the program year as demonstrated by each program's scorecard performance. The report provides a comparison of actual and target results for each program and also provides information in support of the Company's 2015 Demand Side Management Incentive Deferral Account (DSMIDA), Demand Side Management Variance Account (DSMVA), and the Lost Revenue Adjustment Mechanism Variance Account (LRAMVA) claims. Enbridge and Union Gas consulted to align on the general format of the Annual Report and as noted in the Guidelines, the draft version of the report was used by the third party auditor, and updated and finalized by the utilities to reflect the verification of the auditor.

4 Assumes 1.875kg of CO2 are emitted for each m3 gas that is consumed
5 Assuming the average automobile produces 5.1 tonnes of CO_2 per year

2. Demand Side Management Framework

2.1 2015 DSM Plan

On June 30th, 2011, the Board issued DSM Guidelines for the next Multi-Year Plan period titled the "2012 Demand Side Management Guidelines for Natural Gas Utilities" ("2012-2014 DSM Guidelines"). In response, Enbridge undertook an extensive consultation process during the plan development phase and worked with stakeholders on the 2012 DSM Plan budget allocation, scorecards, metrics and targets. On November 4th, 2011, Enbridge submitted its plan outlining proposed DSM activities for the 2012 to 2014 period (EB-2011-0295). On February 28, 2013, and in accordance with the Board Guidelines, Enbridge filed a 2013-2014 DSM Plan Update. Both of the filings were submitted with full Settlement Agreement and were approved by the Board.

On March 31st, 2014 the Minister of Energy issued a Directive to the Board calling for the development of a new DSM policy framework. This new framework was to span a period of six years beginning January 1st, 2015 and, among other things, enable the achievement of all cost-effective DSM.

On September 15th, 2014 the Board issued a Draft Report of the Board outlining its proposed 2015-2020 DSM Framework for Natural Gas Distributors (EB-2014-0134) and called upon all interested parties to provide comment.

On October 15th, 2014 Enbridge, Union Gas, and a wide variety of stakeholders provided comment on the Board's proposed 2015-2020 DSM Framework. An important element of Enbridge's submission was a request that 2015 be treated as a Transition Year, as 2015 is the first year of the 2015-2020 DSM Framework. Among other reasons, Enbridge made this request to satisfy the market's need for certainty and demonstrate that the current DSM consultation process could continue to yield efficient and effective outcomes.

On December 22nd, 2014 the Board released a Report of the Board: Demand Side Management Framework for Natural Gas Distributors (2015-2020) ("Framework") and an accompanying set of Filing Guidelines. Section 15.1 of that Framework set out the Board's direction regarding activities in 2015, calling for 2014 DSM activities to be rolled forward into 2015 in order to facilitate a smooth and measured evolution into the new DSM Framework. Section 15.1 has been included here in its entirety for convenience:

> **15.1 DSM Activities in 2015**
>
> The gas utilities should roll-forward their 2014 DSM plans, including all programs and parameters (i.e., budget, targets, incentive structure) into 2015. Both Enbridge and Union requested that their 2014 activities be rolled-forward into 2015 to help facilitate a smooth evolution into the new DSM framework.
>
> The Board agrees this is appropriate and will allow the gas utilities to fully consider the new DSM framework and appropriately develop their DSM portfolios and suite of programs that will make up their new multi-year plans. The gas utilities should increase their budgets, targets and shareholder incentive amounts in the same manner as they have done throughout the current DSM framework (i.e., 2013 updates to 2014 should now apply as 2014 updates to 2015). The Board expects the gas utilities' new multiyear DSM plans will fully address the guiding principles and key priorities outlined in the framework.
>
> Currently, DSM amounts have already been approved and are included in rates for both Enbridge and Union[25]. If necessary, the gas utilities may modify their current suite of programs and re-allocate funds between approved programs up to a maximum of 30% of the approved annual DSM budget for an individual DSM program. Additionally, the gas utilities may increase overall spending by up to 15%, consistent with the Board's guidance as part of the gas utilities' current, approved DSM plans, and use these additional funds to begin to incorporate and address the guiding principles and key priorities outlined in the DSM framework. If a gas utility incurs DSM spending greater than that which has been previously approved, it should track these expenditures in the DSM variance account for clearance in a future proceeding.
>
> ---
>
> [25] 2015 DSM amounts were approved by the Board as part of EGD's 2014-2018 Custom IR Rate Application (EB2012-0459). EGD has subsequently updated its 2015 DSM budget amounts as part of its 2015 rate application (EB2014-0276). 2015 DSM amounts were approved by the Board as part of Union's 2014-2018 rate application, EB2013-0202. Union has subsequently updated its 2015 DSM budget amounts as part of its 2015 rate application (EB-2014-0271).

With the Framework being issued only a little more than one week prior to the commencement of the 2015 year, it was recognized that appropriate transitional provisions were required to provide the certainty that the gas utilities required in order to be able to effectively operate DSM programs in 2015. Rather than require the utilities to operate their DSM programs in a climate of uncertainty until a decision was issued in that proceeding, the Board ordered a rollover of the 2014 budgets and targets.

Specifically, the Board requested that the gas utilities increase their budgets, targets and shareholder incentive amounts in the same manner as they did to transition from 2013 to 2014.

It should be noted that Enbridge did work extensively with intervenors with a view to attempt to reach an agreement for the purposes of proposing budgets and targets that would have deviated from a strict rollover; however, these discussions did not result in a Settlement that was presented to the Board. As a result, the Company proceeded with its portfolio of DSM program offers relying upon the transitional provisions set out in Section 15.1 of the Framework.

For the purposes of the Update, which the Company filed with the Board for the years 2013 and 2014 (EB-2012-0394) and which was the subject of a complete settlement and acceptance by the Board, a 2% GDP-IPI figure was used to update the budget in both years. Accordingly, the Company updated its 2014 budget by the same 2%[6] consistent with Section 15.1 of the Framework. This 2% change resulted in an increase of the 2014 DSM budget of $32.16 million to a budget of $32.80 million for 2015. Section 15.1 of the Framework also called upon the utilities to increase their shareholder incentives in the same manner as was done for 2013 and 2014. As a result, Enbridge's maximum 2014 shareholder incentive of $10.87 million was increased to a maximum 2015 shareholder incentive of $11.09 million.

Additionally, Section 15.1 allowed the gas utilities to increase overall spending by up to 15% to incorporate the guiding principles and key priorities as outlined in the Framework resulting in an incremental budget of $4.92 million. This 15% incremental budget was incremental to the additional program cost spending previously permitted through the DSMVA.

The Company's 2015 DSM year was delivered consistently with the transitional provisions as set out in the Framework. The Company used the 2014 budget and program targets and escalated these by the rate agreed to by the parties, and accepted by the Board, for the 2013 and 2014 DSM plan years. The Company's activities in 2015 were therefore, based on an expected DSM budget of $32.80 million plus an additional budget of 15% to account for new activities in pursuit of the Board's guiding principles and key priorities of the Multi-Year Plan.

6 EB-2012-0394, Exhibit B, Tab 2, Schedule 9, page 8

In the Decision and Order, published January 20th, 2016, the Board agreed and determined that Enbridge reasonably interpreted the DSM Framework and subsequently approved the Company's 2015 budget, metrics and targets for all scorecards, shareholder incentive amounts, and incremental budget as filed.

2.2 Program and Portfolio Design

Enbridge's 2015 DSM Plan includes three distinct programs; Resource Acquisition, Low Income and Market Transformation. Within each of these programs, Enbridge makes a variety of energy efficiency offers available in support of its customers and the province's GHG emission reduction efforts.

The Resource Acquisition program and its offers focus on achieving direct, measurable savings customer by customer and commonly involve the installation of energy efficient equipment or the implementation of operational improvements. These improvements are often supported by technical assistance and financial incentives among other approaches.

The activities undertaken and offers made available in the Low Income program are largely similar to those included within Resource Acquisition. However, delivering energy efficiency to the low income market presents a unique set of challenges and requires a tailored approach. While the Low Income program will often yield lower net TRC benefits relative to Resource Acquisition, delivery of energy efficiency to these consumers yields various benefits which are difficult to quantify, justifying a Board-approved threshold for cost-effectiveness which is lower than that of Resource Acquisition.

Lastly, Enbridge's Market Transformation program focuses on facilitating fundamental changes in the market, such as increased market shares of energy efficient products and services, or the influencing of consumer behavior and attitudes to reduce the consumption of natural gas. Enbridge's Market Transformation offers have a long-term and holistic view of the use of energy in Ontario and seek to operate where competitive forces are not expected to yield the results sought within an acceptable timeframe.

2.3 Cost-Effectiveness Screening

The utility is expected to assess their DSM portfolio through a method of calculating and testing the cost-effectiveness of its programs. As outlined in the Framework, beginning in 2015, the Board adopted "an enhanced TRC test, or the "TRC-Plus" test, which the gas utilities should use to screen all potential DSM programs when developing their multi-year DSM plans."[7] The utilities were instructed to apply a 15% non-energy benefit adder to the benefit side of the TRC test calculation. Furthermore, the Board directed the utilities to also "incorporate the PAC test as a secondary cost-effectiveness reference tool to help better inform which programs should be proposed."[8]

"The TRC-Plus test measures the benefits and costs of DSM programs for as long as those benefits and costs persist and applies a 15% non-energy benefit adder."[9] The 15% non-energy benefit adder accounts for other benefits not related to the reduction in natural gas such as environmental, economic and social benefits.

In the case of the Resource Acquisition program, if the TRC-Plus ratio (which compares the present value of the natural gas, electricity and water savings and 15% non-energy benefits adder to the present value of the costs) exceeds 1.0, the program is considered cost-effective.

In recognition that the Low Income program may include additional benefits that are not reflected in the TRC-Plus test, the Low Income program is screened using a TRC-Plus threshold of 0.7.

As highlighted in the Guidelines, some programs, such as Market Transformation are not typically amenable to a screening approach (such as TRC-Plus) and instead are reviewed on a case-by-case basis.

The Company has also applied the Program Administrator Cost ("PAC") test as a secondary reference tool in assessing the programs' cost-effectiveness. As outlined in the Guidelines, "the costs included in the PAC test calculation include all expenditures

7 EB-2014-0134. Report of the Board. Demand Side Management Framework for Natural Gas Distributors (2015-2020), OEB, December 22, 2014, Page 33.

8 Ibid, Page 33.

9 EB-2014-0134. Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2020), OEB, December 22, 2014, Page 26.

by the utility to administer DSM programs (i.e., costs to design, plan, administer, deliver, monitor and evaluate)."[10]

The Annual Report provides an opportunity to report both TRC-Plus and PAC assessments for the 2015 DSM program results. Cost-effectiveness screening for 2015 programs is summarized in Section 4.3.

2.4 Program Evaluation

As outlined in the Framework, beginning in 2015, the Board introduced that it would be taking on the coordination function of the EM&V process. Additional clarity regarding the evaluation process was provided on August 21st, 2015, in the memo from the Board to the utilities and to participants in the EB-2014-0134 consultation (EB-2015-0245). The focus of the memo was the establishment of the OEB's process to evaluate the results of Natural Gas Demand Side Management (DSM) programs from 2015 to 2020. This document included the following evaluation responsibilities:

- The Evaluation Contractor would draft an Evaluation, Measurement & Verification (EM&V) Plan for the natural gas utilities' DSM programs for approval by the OEB.
- The Evaluation Advisory Committee (EAC), which includes representation from each of the utilities, would provide advice and input on the development of the plan as required.

The scope of work included in the Request for Proposal posted by the Board on February 8, 2016 for the purpose of selecting an Evaluation Contractor (EC) included further detail explaining the program evaluation process. The utilities' Evaluation Plans that were included in its 2015-2020 Multi-Year DSM Plan, would be reviewed as part of the EC's development of the EM&V Plan and guide the verification tasks, impact assessments and other evaluation studies undertaken in relation to DSM programs.

2.5 Audit of the 2015 DSM Results

The Board's August 21st, 2014 memo (EB-2015-0245) specified that the OEB would be responsible for coordinating and overseeing the evaluation and audit process, including selecting a third party Evaluation Contractor (EC) and publishing the final evaluation

10 Ibid. Page 26

results on an annual basis. The EC will carry out the evaluation and audit processes of all DSM programs and provide an opinion on whether the claimed DSM Incentive (DSMI) amount, Lost Revenue Adjustment Mechanism Variance Account (LRAMVA), and Demand Side Management Variance Account (DSMVA) have been correctly calculated using reasonable assumptions. The Evaluation Advisory Committee (EAC) which includes utility representation as described in Section 2.6 will provide input and play an advisory role throughout the audit to facilitate the achievement of the audit objectives.

2.6 Evaluation Advisory Committee (EAC)

As detailed in the August 21st, 2015 memo from the Board (EB-2015-0245), the EAC provides input and advice as required throughout the DSM evaluation process. The EAC is comprised of:

- Experts representing non-utility stakeholders, with demonstrated experience and expertise in the evaluation of DSM technologies and programs, natural gas energy efficiency technologies, multi-year impact assessments, net-to-gross studies, free ridership analysis and natural gas energy efficiency persistence analysis;
- Expert(s) retained by the OEB;
- Representatives from the Independent Electricity System Operator (IESO);
- Representatives from each natural gas utility; and
- Representatives from the Ministry of Energy (MOE) and the Environmental Commissioner of Ontario (ECO), who will participate as observers.

The OEB has appointed the following non-utility stakeholders as members of the EAC:

- Chris Neme, Energy Futures Group
- Jay Shepherd, Jay Shepherd Professional Corporation
- Marion Fraser, Fraser & Company

Non-utility stakeholders are expected to provide input and advice based on their experience and technical expertise and not to advocate positions of parties they have represented before the OEB in various proceedings.

2.7 Transition Plan of TEC Activities to the OEB

As outlined in the letter from the Board dated March 4th, 2016 (EB-2015-0245), the TEC's evaluation activities will be transitioned to the OEB under the new DSM evaluation governance structure. Further discussion with OEB Staff and the TEC has provided additional clarity/direction on the following specific projects:

Technical Reference Manual (TRM) Development. Development of the TRM with updated measures and input assumptions is near completion and the TEC will continue to finalize the TRM. The management of the online portion of the TRM has been transitioned to OEB Staff, who will post the final TRM online when it is available. The utilities will continue to manage any remaining contractual obligations and payments related to the TRM.

Custom Project Net-to-Gross Study. Following input from the TEC on a draft work plan prepared by the project consultant currently under contract, this study will be transitioned to OEB Staff. The utilities will continue to manage contractual obligations and payments associated with this project. OEB Staff will assume oversight of the study with input from the EAC, and will confirm the completion of major milestones for the utilities to process payments of consultant's invoices.

Boiler Baseline Study. The TEC will select the Boiler Baseline proponent with input from Board Staff. This will be the last order of business for the TEC on this project. The utilities will take over administrative responsibility and accountability for the study following selection of proponent. The EAC will provide input to the utilities on the study as appropriate.

Persistence Study. OEB Staff will be responsible for the procurement process and management of the Persistence Study, including management of project deliverables and contractual obligations through to completion of the study, with input from the EAC.

3. OEB Data Reporting Requirements

The following tables summarize the annual reporting key elements outlined in Section 14.2 of the Guidelines.

Table 3.0 Annual and Long-Term DSM Budgets

($/year and $/6 years)

OEB Approved Budgets							
Resource Acquisition (RA)	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**Total**
Residential	$1,872,720	$13,024,688	$16,705,000	$20,175,000	$20,578,500	$20,990,070	$93,345,978
Commercial / Industrial	$12,571,070	$16,278,937	$17,679,381	$17,737,977	$16,355,713	$16,685,480	$97,308,558
RA Program Costs	$14,443,790	$29,303,625	$34,384,381	$37,912,977	$36,934,213	$37,675,550	$190,654,536
RA Overheads	$4,731,485	$5,033,048	$5,104,327	$5,249,479	$5,122,057	$5,232,967	$30,473,363
Total RA	**$19,175,275**	**$34,336,673**	**$39,488,708**	**$43,162,456**	**$42,056,270**	**$42,908,517**	**$411,782,435**
Low Income (LI)							
LI Program Costs	$6,864,090	$10,201,788	$10,908,121	$11,690,496	$11,923,306	$12,160,772	$63,748,573
LI Overheads	$517,988	$1,743,622	$1,619,299	$1,618,681	$1,653,531	$1,689,078	$8,842,199
Total LI	**$7,382,078**	**$11,945,410**	**$12,527,420**	**$13,309,177**	**$13,576,837**	**$13,849,850**	**$72,590,772**
Martket Transformation (MT)							
MT Program Costs	$4,890,900	$5,614,683	$5,849,381	$6,045,400	$6,174,079	$6,305,335	$34,879,778
MT Overheads	$1,353,687	$964,351	$868,335	$837,054	$856,225	$875,783	$5,755,435
Total MT	**$6,244,587**	**$6,579,034**	**$6,717,716**	**$6,882,454**	**$7,030,304**	**$7,181,118**	**$40,635,213**
Total Program Costs (without overheads)	$26,198,780	$45,120,096	$51,141,883	$55,648,873	$55,031,598	$56,141,657	$289,282,887
Total Program Overheads	$6,603,160	$7,741,021	$7,591,961	$7,705,214	$7,631,813	$7,797,828	$45,070,997
Total Program Costs (with overheads)	**$32,801,939**	**$52,861,117**	**$58,733,844**	**$63,354,087**	**$62,663,411**	**$63,939,485**	**$334,353,883**
Portfolio Overheads							
EM&V	*n/a*	$1,500,000	$1,700,000	$1,700,000	$1,736,746	$1,774,228	$8,410,974
Collaboration & Innovation	*$1,000,000*	$1,000,000	$1,000,000	$1,000,000	$1,021,616	$1,043,663	$6,065,279
DSM IT	*n/a*	$1,000,000	$1,000,000	$1,000,000	$1,000,000	$1,000,000	$5,000,000
Energy Literacy	*n/a*	$0	$500,000	$500,000	$0	$0	$1,000,000
Total Portfolio Overheads	*n/a*	$3,500,000	$4,200,000	$4,200,000	$3,758,362	$3,817,891	$19,476,253
2015 Incremental Budget	$4,920,291	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*
Total Portfolio Budget	**$37,722,230**	**$56,361,117**	**$62,933,844**	**$67,554,087**	**$66,421,773**	**$67,757,376**	**$358,750,427**

1. In 2015 the Collaboration and Innovation amount of $1M was included in the Incremental budget of $4.92M

2. Total Collaboration & Innovation budget as approved by the Board is $6M for 2015-2020

Table 3.1 Actual Annual Total DSM Costs

(including DSM spending[11], overheads, evaluation, shareholder incentive, lost revenues) for each rate class dating back to 2007

Annual Actual Total DSM Costs									
RATE CLASS	2007	2008	2009	2010	2011	2012	2013	2014	2015
RATE 1	$11,894,135	$12,545,981	$14,794,795	$12,467,796	$14,214,627	$17,935,484	$13,881,901	$23,507,037	$26,855,974
RATE 6	$2,848,384	$7,519,262	$7,486,577	$10,713,308	$15,103,141	$17,127,050	$15,172,590	$13,901,251	$15,646,361
RATE 9	$0	$0	$0	$0	$0	$1,425	$1,420	$1,712	$1,839
RATE 100	$8,949,764	$3,201,527	$2,667,170	$86,297	$17,677	$0	$0	$0	$0
RATE 110	$3,658,449	$1,041,758	$1,943,819	$1,470,858	$1,048,222	$783,904	$937,258	$1,189,687	$1,904,974
RATE 115	$643,144	$1,716,735	$1,314,146	$545,382	$602,386	$1,329,072	$1,420,390	$567,271	$662,208
RATE 125	$0	$0	$0	$0	$0	$53,449	$53,268	$64,223	$68,967
RATE 135	$1,762	$79,757	$11,685	$59,163	$121,756	$441,318	$320,401	$123,739	$59,072
RATE 145	$855,487	$901,590	$676,730	$729,534	$655,237	$495,925	$369,074	$253,864	$153,885
RATE 170	$294,508	$1,860,562	$1,843,628	$2,040,735	$2,195,089	$536,445	$149,399	$457,841	$403,612
RATE 200	$0	$0	$0	$0	$0	$18,529	$18,466	$22,264	$23,909
RATE 300	$0	$0	$0	$0	$0	$3,563	$3,551	$4,281	$4,598
TOTAL	$29,145,632	$28,867,172	$30,738,550	$28,113,075	$33,958,134	$38,726,165	$32,327,718	$40,093,170	$45,785,399

Table 3.2 Historic Actual Annual DSM Spending

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Total DSM Spending ($ millions) [1]	$21.20	$23.03	$25.42	$24.00	$27.24	$30.61	$27.84	$32.51	$35.78

1. Total DSM Spending includes variable costs, fixed costs and DSMVA where applicable

Table 3.3 DSM Spending as a Percent (%) of Distribution Revenue

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Total DSM Spending (millions $) [1]	$21.2	$23.0	$25.4	$24.0	$27.2	$30.6	$27.8	$32.5	$35.8
Total Distribution Revenue (millions $) [2 3 4 5]	$980.9	$995.9	$1,012.1	$960.4	$978.8	$972.0	$1,055.0	$1,044.0	$1,055.4
DSM Spending as % of Distribution Revenue	2.2%	2.3%	2.5%	2.5%	2.8%	3.1%	2.6%	3.1%	3.4%

1. Total DSM Spending includes variable costs, fixed costs and DSMVA where applicable
2. Distribution Revenue is equal to the gas distribution margin, and is the gas sales and distribution revenue less the cost of gas
3. Distribution Revenue includes gas sales and transportation of gas less gas commodity cost
4. Distribution Revenue excludes transmission, compression, and storage
5. Distribution Revenue is based on data unnormalized for weather

11 As the request is for actual costs, Enbridge interprets this to be 'DSM spending' rather than 'DSM budget' as written in Section 14.2 of the Guidelines.

Table 3.4 Historic Shareholder Incentive Amounts Available and Earned

	2007	2008	2009	2010	2011	2012 [1]	2013	2014	2015 [2]
	$ millions								
Total Shareholder Incentive Earned	$8.25	$5.80	$5.36	$4.16	$6.77	$8.16	$4.54	$7.65	$10.08
Maximum Shareholder Incentive Available	$9.00	$9.22	$9.24	$9.40	$10.16	$10.45	$10.66	$10.87	$11.09

1. 2012 Shareholder Incentive includes reduction of -$657,223 per Board's decision (EB-2013-0352)
2. 2015 Shareholder Incentive subject to Board approval

Table 3.5 Shareholder Incentive Earned as a Percent (%) of DSM Spending[12]

	2007	2008	2009	2010	2011	2012 [2]	2013	2014	2015 [3 4]
Total Shareholder Incentive ($ million)	$8.25	$5.80	$5.36	$4.16	$6.77	$8.16	$4.54	$7.65	$10.08
Total DSM Spending [1]	$21.20	$23.03	$25.42	$24.00	$27.24	$30.61	$27.84	$32.51	$35.78
Total DSM Spending as a % of Shareholder Incentive Earned	39%	25%	21%	17%	25%	27%	16%	24%	28%

1. DSM spending includes variable costs, fixed costs, and overheads
2. 2012 Shareholder Incentive includes reduction of -$657,223 per Board's decision (EB-2013-0352)
3. 2015 Shareholder Incentive subject to Board approval
4. 2015 DSM Spending includes incremental spending of $559,378

Table 3.6 Annual and Long-Term Natural Gas Savings Targets

	Annual Natural Gas Savings Targets					
Scorecard	2015	2016	2017	2018	2019	2020
Resource Acquisition (m^3)	**1,011,901,200**	**983,790,685**	*Targets are formulaic based on past year's performance*			
Low-Income (m^3)	**92,800,000**	**96,690,000**				

12 Enbridge interprets this request as requesting values as a percentage of 'DSM spending' rather than 'DSM budget' as written in Section 14.2 of the Guidelines.

Table 3.7 2015 Total Annual & Cumulative Natural Gas Savings (Gross and Net)

	2015 Annual Gas Savings [1]		2015 Cumulative Gas Savings [1]	
	Gross	Net	Gross	Net
Resource Acquisition	62,780,541	44,698,972	1,021,749,160	734,128,834
Low-Income	4,306,970	4,272,585	92,380,469	92,036,617
Total	**67,087,511**	**48,971,556**	**1,114,129,629**	**826,165,451**

1. 2015 DSM results subject to Board approval

Table 3.8 Total Historic Annual Natural Gas Savings (Gross and Net)

	2007	2008	2009	2010	2011	2012	2013	2014	2015 [1]
Total Net Gas Savings (millions m3)	85.07	77.25	69.86	64.58	76.40	60.14	47.74	43.54	48.97
Total Gross Gas Savings (millions m3)	85.99	121.98	117.62	98.82	114.14	92.53	66.06	60.62	67.09

1. 2015 DSM results subject to Board approval

Table 3.9 Total Historic Cumulative Natural Gas Savings (Gross and Net)

	2007	2008	2009	2010	2011	2012	2013	2014	2015 [1]
Total Net CCM (millions m3)	1,214.10	1,118.98	1,039.18	951.40	1,253.82	1,068.98	826.91	719.84	826.17
Total Gross CCM (millions m3)	1,233.54	1,809.65	1,801.77	1,455.74	1,811.35	1,593.05	1,148.12	993.62	1,114.13

1. 2015 DSM results subject to Board approval

Table 3.10 Total Annual Natural Gas Savings as Percent (%) of Total Annual Natural Gas Sales

(Gross and Net)

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Net Gas Savings Total (millions m3) [1]	85.1	77.3	69.9	64.6	76.4	60.1	47.7	43.5	49.0
Net Gas Savings as % of Total Gas Sales	0.7%	0.7%	0.6%	0.6%	0.7%	0.6%	0.4%	0.4%	0.4%
Gross Gas Savings Total (millions m3) [1]	86.0	122.0	117.6	98.8	114.1	92.5	66.1	60.6	67.1
Gross Gas Savings as % of Total Gas Sales	0.7%	1.0%	1.1%	0.9%	1.0%	0.9%	0.6%	0.5%	0.6%
Consumption (millions m3) [2]	11,862.9	11,686.5	11,114.9	10,742.3	11,303.2	10,304.4	11,338.3	12,434.3	11,728.3

1. 2015 DSM results are subject to Board approval

2. Annual consumption volumes include rate classes that are subject to DSM costs only. Rates 9, 125, 200 and 300 are excluded as they do not participate in DSM

Table 3.11 Total Cumulative Natural Gas Savings as Percent (%) of Total Annual Natural Gas Sales

(Gross and Net)

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Net CCM Total (millions m3) [1]	1,214.1	1,119.0	1,039.2	951.4	1,253.8	1,069.0	826.9	719.8	826.2
Net Gas Savings as % of Total Gas Sales	10.2%	9.6%	9.3%	8.9%	11.1%	10.4%	7.3%	5.8%	7.0%
Gross CCM Total (millions m3) [1]	1,233.5	1,809.7	1,801.8	1,455.7	1,811.3	1,593.0	1,148.1	993.6	1,114.1
Gross Gas Savings as % of Total Gas Sales	10.4%	15.5%	16.2%	13.6%	16.0%	15.5%	10.1%	8.0%	9.5%
Consumption (millions m3) [2]	11,862.9	11,686.5	11,114.9	10,742.3	11,303.2	10,304.4	11,338.3	12,434.3	11,728.3

1. 2015 DSM results are subject to Board approval

2. Annual consumption volumes include rate classes that are subject to DSM costs only. Rates 9, 125, 200 and 300 are excluded as they do not participate in DSM

Table 3.12 Actual Annual Gas Operating Revenue

	2007	2008	2009	2010	2011	2012	2013	2014	2015
Total Operating Revenue (millions $) [1]	**$3,095.0**	**$3,233.8**	**$2,952.3**	**$2,394.1**	**$2,393.6**	**$2,240.9**	**$2,613.4**	**$2,861.3**	**$2,892.1**
Less Total Gas Cost (millions $) [2]	**$2,113.0**	**$2,236.1**	**$1,938.6**	**$1,432.3**	**$1,413.3**	**$1,267.6**	**$1,556.8**	**$1,815.5**	**$1,834.8**
Total Distribution Revenue (millions $) [3]	**$982.0**	**$997.7**	**$1,013.7**	**$961.8**	**$980.3**	**$973.3**	**$1,056.6**	**$1,045.8**	**$1,057.3**

1. Operating Revenue includes gas sales and transportation, transmission, compression, and storage. All values are unnormalized for weather

2. Gas Cost is based on data unnormalized for weather

3. Distribution revenue is equal to the gas distribution margin and is the gas sales plus transportation less the cost of gas

Table 3.13 Total Natural Gas Sales per Rate Class Subject to DSM Costs

Rate Class	2015 Natural Gas Volumes (millions m3)
General Service	
Rate 1	4,997.0
Rate 6	5,006.6
Total General Service	**10,003.6**
Contract Service	
Rate 100	3.7
Rate 110	667.9
Rate 115	512.2
Rate 135	68.6
Rate 145	77.5
Rate 170	394.8
Total Contract Service	**1,724.7**
Grand Total	**11,728.3**

**Natural Gas Sales (Volumes) for rate classes that are subject to DSM only*

Table 3.14 Number of Customers by Customer Type

Customer Type	# of Customers 2015
Residential [1]	1,930,657
Commercial	157,762
Industrial	6,262
Total	**2,094,681**

1. Residential customers include Low Income, which cannot be differentiated

Table 3.15 Number of Customers Broken Out by Rate Class

Rate Class	# of Customers 2015
Rate 1	1,930,657
Rate 6	163,634
Rate 9	6
Rate 100	2
Rate 110	227
Rate 115	25
Rate 125	5
Rate 135	42
Rate 145	52
Rate 170	26
Rate 200	1
Rate 300	2
Rate 315	2
Total	**2,094,681**

4. 2015 DSM Program Results Summary

4.1 2015 DSM Scorecard Summary

The 2015 DSM program scorecard performance is presented in Table 4.0.

Table 4.0 2015 DSM Program Scorecard Summary

	Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Results
Resource Acquisition	Volumes	*Cumulative Savings (million m^3)*	**92%**	758.9	1,011.9	1,264.9	**734.13**
Resource Acquisition	Residential Deep Savings	*Number of Participants* [1]	**8%**	571	762	952	**5,646**
Low Income	Single Family (Part 9)	*Cumulative Savings (million m^3)*	**50%**	18.1	24.1	30.2	**28.07**
Low Income	Multi-residential (Part 3)	*Cumulative Savings (million m^3)*	**45%**	51.6	68.7	86.0	**63.97**
Low Income	Multi-residential (Part 3) LIBPM [2]	*% of Part 3 Participants Enrolled* [3]	**5%**	30%	40%	50%	**65%**
Market Transformation	Residential Savings by Design	*Completed Units*	**40%**	833	1,111	1,389	**1,987**
Market Transformation	Residential Savings by Design	*Builders Enrolled* [4]	**60%**	13	18	22	**19**
Market Transformation	Commercial Savings by Design	*New Developments Enrolled*	**100%**	11	18	24	**24**
Market Transformation	Home Labelling	*Realtor Commitments* [5]	**50%**	N/A	5,001 [6]	10,001 [6]	**41,650**
Market Transformation	Home Labelling	*Ratings performed by buyers and/or sellers*	**50%**	2,250	4,500	6,750	**333**

1. Number of participants (houses) with at least two major measures and where average annual gas savings across all participants is at least 25% of combined baseline usage.

2. LIBPM - Low Income Building Performance Management is the Low Income offer complement to the Commercial Run It Right (RIR) offer.

3. Low Income Building Performance Management (LIBPM) percentage of Part 3 buildings enrolled in the current year program = (x+y)/(x+y+z):

   x = # of new LIBPM buildings in the current year that have participated in another aspect of the Low Income program in a previous year of 2012-2014 plan; y = # of new LIBPM buildings participating in current year that have not previously participated in the Low Income program; z = # of buildings in the current year that have implemented custom projects other than LIBPM.

4. Eligible builders based on a minimum of 50 homes built in the prior year.

5. Commitments to make provision for a data field to show home energy ratings for all homes listed by participating realtors (industry-wide commitment to include such a field on MLS or similar listing service and/or realtors' commitment to do so with all the homes they list on their own websites, handouts and other consumer material).

6. Commitment from realtors collectively responsible for more than 5,000 (middle target) or 10,000 (upper target) listings/year.

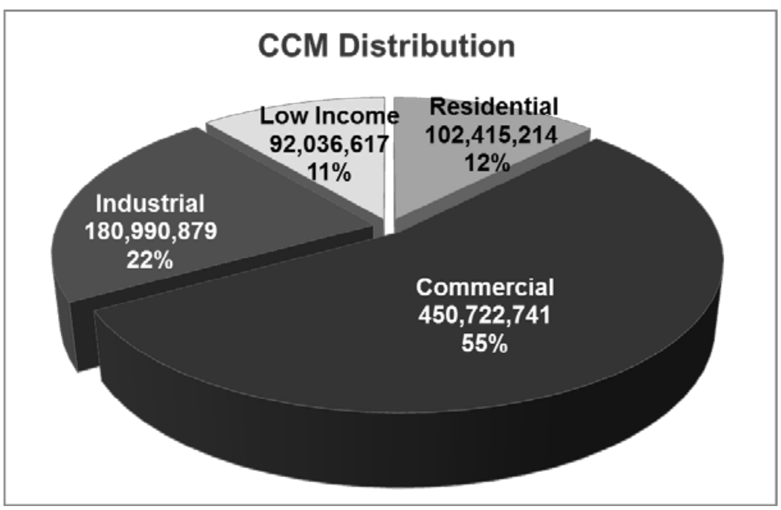
The 2015 weighted scorecard is the basis for the calculation of the Demand Side Management Shareholder Incentive. DSMI amounts for the 2015 program year are outlined in Section 9 of this report.

Table 4.1 2015 CCM Savings Results by Sector

Program/Sector	2015 Net CCM Results (m^3)
Resource Aquisition	
Residential	*102,415,214*
Commercial	*450,722,741*
Industrial	*180,990,879*
Resource Acquisition Total	**734,128,834**
Low Income	**92,036,617**
Combined Total	**826,165,451**

As summarized in Table 4.1, in terms of Net CCM savings, 2015 results totalled 826,165,451 cumulative m^3 for all offers that include CCM as a metric. In 2015, the Commercial sector was the largest overall contributor to CCM savings, accounting for 450,722,741 CCM or 55% of the total net CCM results. Industrial sector offers contributed 22% of the total CCM savings followed by the Residential sector and the Low Income program responsible for 12% and 11% of CCM, respectively.

Table 4.2 Distribution of 2015 Net CCM Results



In 2015, Enbridge delivered three Market Transformation offers, all of which performed well in relation to performance targets. On a weighted scorecard basis, all three offers met or exceeded upper targets outlined in the scorecard. Results for the Market Transformation program offers are reviewed in Section 7 of this report.

4.2 Annual and Cumulative (Gross and Net) Results

As outlined in the Guidelines, the utilities "should provide the annual and cumulative resource savings attributable to each program, presented as both net and gross of the adjustment factors."[13]

Table 4.3 2015 Annual and Cumulative Natural Gas Savings

	Program/Sector/Offer	Gross Annual Gas Savings (m^3)	Net Annual Gas Savings (m^3)	Gross CCM (m^3)	Net CCM (m^3)
Resource Acquisition	**Residential**				
	Home Energy Conservation	7,956,225	6,762,791	120,488,487	102,415,214
	Total Residential	**7,956,225**	**6,762,791**	**120,488,487**	**102,415,214**
	Commercial				
	Commercial Custom	23,293,072	19,434,966	426,012,146	350,622,209
	Commercial Prescriptive	6,774,554	5,674,927	116,504,921	97,416,428
	Run It Right	536,821	536,821	2,684,105	2,684,105
	Total Commercial	**30,604,447**	**25,646,715**	**545,201,172**	**450,722,741**
	Industrial				
	Industrial Custom	23,658,347	11,890,642	345,232,715	173,397,871
	Industrial Prescriptive	561,521	398,824	10,826,785	7,593,008
	Total Industrial	**24,219,869**	**12,289,466**	**356,059,500**	**180,990,879**
Low Income	**Low Income**				
	Single Family (Part 9)	1,135,609	1,129,070	28,132,657	28,067,264
	Multi-Residential (Part 3)	3,171,361	3,143,515	64,247,812	63,969,353
	Total Low Income	**4,306,970**	**4,272,585**	**92,380,469**	**92,036,617**
	Grand Total	**67,087,511**	**48,971,556**	**1,114,129,629**	**826,165,451**

13 EB-2014-0134. Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2020), OEB, December 22, 2014, Page 18 .

Table 4.3 details the annual gas savings and cumulative lifetime natural gas savings results (in cubic meters) for each of the program components that have CCM as a performance metric. Savings results are summarized for both gross and net savings (net of applicable adjustment factors).

4.3 2015 Program Cost-Effectiveness Screening

Table 4.4 summarizes the TRC-Plus screening calculations for the 2015 Enbridge DSM Portfolio for illustrative purposes. The portfolio as a whole was cost-effective with an overall TRC-Plus ratio of 2.95. Further, the two programs to which this screening applies, Resource Acquisition (3.12 TRC-Plus Ratio) and Low Income (1.88 TRC-Plus Ratio) were also cost-effective to deliver as individual programs.

Table 4.4 2015 TRC-Plus Screening Summary

Sector/Offer	NPV TRC Plus Benefits	Total TRC Costs	Net TRC Plus Benefits	TRC Plus Ratio
Residential				
Home Energy Conservation	19,724,692	8,812,469	10,912,223	2.24
All Residential Total	**19,724,692**	**8,812,469**	**10,912,223**	**2.24**
Commercial				
Commercial Custom	86,028,559	26,433,133	59,595,426	3.25
Commercial Prescriptive	20,777,419	3,640,863	17,136,556	5.71
Run It Right	523,655	1,591,860	-1,068,205	0.33
All Commercial	**107,329,633**	**31,665,856**	**75,663,777**	**3.39**
Industrial				
Industrial Custom	32,178,775	5,179,632	26,999,143	6.21
Industrial Prescriptive	1,530,447	304,562	1,225,885	5.03
All Industrial	**33,709,221**	**5,484,194**	**28,225,028**	**6.15**
Overheads		5,639,080	-5,639,080	
Overall Resource Acquisition	**160,763,547**	**51,601,598**	**109,161,948**	**3.12**
Low Income				
Single Family (Part 9)	4,309,787	4,077,752	232,036	1.06
Multi-Residential (Part 3)	10,978,840	3,426,793	7,552,047	3.20
Overheads		617,349	-617,349	
Overall Low Income	**15,288,628**	**8,121,893**	**7,166,734**	**1.88**
Combined RA/Low Income *	**176,052,174**	**59,723,491**	**116,328,683**	**2.95**

*This summary does not include calcuations for the Market Transformation program.

As proposed in the Guidelines, the Company is expected to use the Program Administrator Cost (PAC) test as a secondary reference tool in assessing the programs' cost-effectiveness. Table 4.5 below summarizes the PAC screening calculations for the 2015 Enbridge DSM Portfolio. The portfolio as a whole was cost-effective with an overall PAC ratio of 4.47.

Table 4.5 2015 PAC Screening Summary

Sector/Offer	NPV PAC Benefits	Total PAC Costs	Net PAC Benefit	PAC Ratio
Residential				
Home Energy Conservation	18,205,398	9,362,295	8,843,103	1.94
All Residential Total	**18,205,398**	**9,362,295**	**8,843,103**	**1.94**
Commercial				
Commercial Custom	56,548,553	4,577,666	51,970,887	12.35
Commercial Prescriptive	16,207,128	759,387	15,447,741	21.34
Run It Right	523,655	1,458,896	-935,241	0.36
All Commercial	**73,279,336**	**6,795,949**	**66,483,387**	**10.78**
Industrial				
Industrial Custom	32,084,579	2,139,556	29,945,023	15.00
Industrial Prescriptive	1,385,792	27,150	1,358,642	51.04
All Industrial	**33,470,371**	**2,166,706**	**31,303,665**	**15.45**
Overheads		5,639,080	-5,639,080	
Overall Resource Acquisition	**124,955,105**	**23,964,031**	**100,991,074**	**5.21**
Low Income				
Single Family (Part 9)	4,288,990	4,444,616	-155,626	0.96
Multi-Residential (Part 3)	10,058,221	2,111,746	7,946,475	4.76
Overheads		617,349	-617,349	
Overall Low Income	**14,347,211**	**7,173,710**	**7,173,501**	**2.00**
Combined RA/Low Income *	**139,302,316**	**31,137,741**	**108,164,575**	**4.47**

*This summary does not include calcuations for the Market Transformation program.

5. Resource Acquisition Scorecard

There are two performance metrics in Enbridge's Resource Acquisition scorecard encompassing results attributable to offers which are geared to the Residential, Commercial and Industrial market segments. Performance for the Resource Acquisition program is measured primarily in terms of net CCM of natural gas savings but also includes a residential deep savings metric.

Resource Acquisition offers focus on achieving direct, measureable savings customer by customer and commonly involve the installation of energy efficient equipment or the implementation of operational improvements.

In the residential sector, the Home Energy Conservation (HEC) offer comprises upgrades to space and water heating equipment and home building envelope upgrades. The deep savings metric measures the number of participants in HEC that achieve an average annual gas savings across all participants of at least 25% of combined baseline usage.

For commercial customers, prescriptive and custom project offers are available for new and existing commercial building customers and include the installation of efficient heating, ventilating and air conditioning (HVAC) systems, and custom solutions specific to the customers' needs.

Industrial customers tend to have differing and unique considerations. In addition to selected prescriptive measures, projects for industrial customers are most often customized solutions, engineered to meet the specific needs of a customer's manufacturing process and facility.

Enbridge works across the entire marketplace to build awareness of the energy efficiency opportunities supported through its program. The ongoing education, customer support and technical assistance provided by DSM consultants continue to be a key driver in delivering results for the Resource Acquisition program.

Results for CCM (natural gas savings volumes) in Enbridge's 2015 Resource Acquisition (RA) program were 734.1 million CCM. The Resource Acquisition program scorecard also includes a deep savings metric specific to the Residential sector. There were 5,646 participants counted towards this metric. This result exceeded the upper scorecard target.

Table 5.0 2015 Resource Acquisition Scorecard

Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Volumes	*Cumulative Savings (million m³)*	***92%***	758.9	1,011.9	1,264.9	**734.13**
Residential Deep Savings	*Number of Participants* [1]	***8%***	571	762	952	**5,646**

1. Number of participants (houses) with at least two major measures (average annual gas savings across all participants must be at least 25% of combined baseline usage).

Within the RA program, each of the Residential, Commercial and Industrial sectors contributed to the CCM savings target as detailed below in Table 5.1. Further detail on the offers within each of these sectors is provided in the following pages.

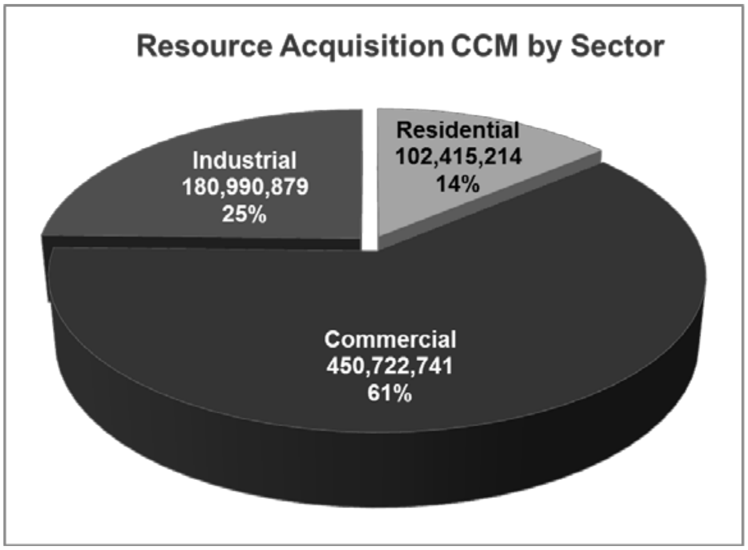
Table 5.1 2015 Resource Acquisition Program Sector Results

Resource Aquisition Program Sector	CCM 100% Target (m^3)	2015 Net CCM (m3)	TRC-Plus Ratio	PAC Ratio	# of Projects [1]	# of Units [2]
Residential	12,024,643	102,415,214	2.24	1.94	5,646	---
Commercial	644,789,155	450,722,741	3.39	10.78	590	16,877
Industrial	355,087,402	180,990,879	6.15	15.45	115	235
Total Resource Acquisition	**1,011,901,200**	**734,128,834**	**3.12**	**5.21**	**6,351**	**17,112**

1. # of Projects summarizes the number of unique projects for custom offers, RIR and HEC.
2. # of Units summarizes the number of units installed for prescriptive offers.

CCM savings contributions from each sector within the RA program are illustrated in Table 5.2. Commercial offers were responsible for 61% of the total CCM savings in the RA program. Industrial and Residential offers contributed 25% and 14% of results, respectively.

Table 5.2 2015 Resource Acquisition – CCM Results by Sector



All Resource Acquistion offers delivered to Enbridge customers in 2015 and discussed below will be continued in the Resource Acquisition DSM program in 2016.

5.1 Residential Resource Acquisition

Home Energy Conservation (HEC)

Objectives	The Residential component of the RA program focuses on the existing home sector through the marketing and delivery of a home energy conservation initiative. The goal of the HEC offer is to achieve deep energy savings in existing homes and to raise awareness of the benefits of energy efficiency. The initiative is designed to reduce gas use for space and water heating using a holistic approach, encouraging conservation through the installation of high efficiency equipment as well as thermal envelope improvements to reduce the space heating load. With financial incentives, the offer helps homeowners make their homes more energy efficient and reduces the burden of high energy costs.
Target Customer	HEC is targeted to Rate 1 residential customers.

Metrics	The first metric is cumulative cubic meter (CCM) savings generated by participants. The second metric is total number of participants – specifically, the number of houses with at least two eligible measures implemented and where average annual gas savings across all participants is at least 25% of baseline usage.
Tracking Methodology	Gas savings are claimed based on results calculated through the use of NRCANs accredited modeling software (HOT2000) utilized by Certified Energy Auditors (CEAs). Reporting provided to the Company by the delivery agents summarizes information regarding participants, dates, measures installed and gas savings (m^3) which are maintained and tracked monthly. The number of participants (houses) with at least two major measures, and where average annual gas savings across all participants is at least 25% of combined baseline usage, are calculated, tracked and counted toward the deep savings participant metric.
Offer Description	Since the cancellation of the federal government funded ecoENERGY program that ran from 2007 and ended in early 2012, there has been a market need for initiatives that drive energy efficiency in the existing housing sector. This offer was introduced midway through 2012 to encourage and support gas savings opportunities in existing residential houses and to meet the priorities outlined in the Board's 2012-2014 DSM Guidelines, in particular, the goal of pursuing deep savings. HEC is designed to capture deep energy efficiency savings opportunities through the delivery of a holistic, "whole home" approach. The HEC offer utilizes accredited software such as Natural Resources Canada's (NRCan) as the foundation in calculating annual gas

	savings for each participant. The software provides an effective building energy simulation tool to model the savings. Participants receive a pre-retrofit energy audit conducted by a certified energy advisor before starting work and a post-retrofit energy audit to calculate gas savings. Measures include home envelope improvements and mechanical system upgrades as these measures offer the greatest opportunity for “deep”, long-term energy conservation through gas savings. Enbridge offers qualifying customers incentive dollars towards the pre-retrofit energy audit of their home and the opportunity for additional incentives if the participant completes at least two upgrades from a list of qualifying measures. The offer aims to ensure that the installation of these measures contributes to the achievement of an average 25% annual gas savings over the participant portfolio, based on pre- and post-energy audit results. The qualifying measures included for HEC are as follows: • Heating system replacement; • Foundation insulation; • Water heating system replacement; • Air sealing; • Attic insulation; • Window replacements; • Wall insulation; • Drain water heat recovery; and • Exposed floor insulation. To be eligible for the offer, customers must meet the following criteria: • Be a residential homeowner in the Enbridge franchise area; • Have a valid Enbridge Gas account in good standing; • The home’s primary heat source must be natural gas; • Use an approved Certified Energy Evaluator/Auditor; • Install at least two measures; and • Complete a pre- and post-energy audit.

	In 2015, to help offset the costs of recommended upgrades, customer incentives of up to $1,600 were available for achieving 25%-49% in annual gas savings and up to $2,000 for achieving 50% and above in annual gas savings.
Cost-Effectiveness	The HEC offer is cost-effective as supported by the TRC-Plus and PAC ratios summarized in Table 5.3 below.
2015 Results	Also as outlined in Table 5.3 below, the HEC offer contributed 102.4 million CCM to the Resource Acquisition results in 2015 with a total of 5,646 participants. These participants counted toward the Residential Deep Savings metric, exceeding the upper target of 952 participants. As communicated in 2015, including during the Oral Hearing for the 2015-2020 Multi-Year DSM Plan proceeding (EB-2015-0049), based on forecast participation in the offer, Enbridge halted the offer mid-year in order to manage budget requirements.

Table 5.3 2015 Residential Resource Acquisition Results

Resource Acquisition Residential Sector	2015 Net CCM (m3)	TRC-Plus Ratio	PAC Ratio	# of Projects
Home Energy Conservation	102,415,214	2.24	1.94	5,646

2015 Commentary and Lessons Learned:

- The HEC offer again demonstrated great success in 2015. A key focus was the continuing expansion of the offer to a broader customer base, working toward the goal of making the offer accessible across the Enbridge franchise area.

- Enbridge continued to provide training sessions and touchpoint meetings to ensure that procedures and processes required for tracking were understood and followed.

- Enbridge continue to work in 2015 with the City of Toronto on the Home Energy Loan Program (HELP) to further broaden the delivery of the HEC offer in Toronto with a simultaneous expansion of the regions that could qualify for HELP. The initiative provides a financing tool offered by the City to assist homeowners with improving their home's energy efficiency and save money. Low interest loans are available to qualifying homeowners with repayment facilitated through installments on property tax bills.

- Marketing efforts for HEC have been well received and included:
  - Enbridge Channel Consultants marketing to Heating, Ventilation and Air Conditioning (HVAC) and insulation contractors through e-blasts to communicate updates, geographic expansion and to promote opportunities for residential customers;
  - enhancements to Enbridge's residential energy efficiency microsite - www.knowyourenergyscore.ca in order to increase user-friendliness;
  - targeted advertising in lifestyle magazines to highlight the HEC offer and gas savings opportunities directly to homeowners;
  - participation and exhibition at franchise area home shows to promote the HEC offer and increase awareness;
  - collaboration with the Toronto and Region Conservation Authority (TRCA) Sustainable Neighbourhood Action Planning (SNAP) for delivery of HEC

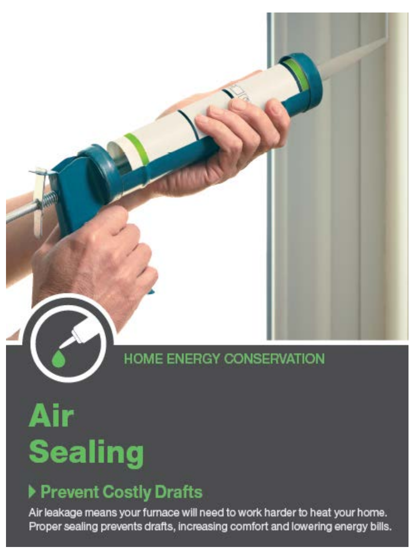
marketing information to increase consumer awareness and drive participation; and,

- promotion to the realtor community in addition to the Home Labelling communication and marketing efforts for that initiative.

➢ Overall, net annual gas savings per project averaged approximately 1,200 m^3. The majority of participants implemented heating system replacements; the next most common measures installed were air sealing and attic insulation.

➢ To support the growing momentum of this offer in 2015 and the opportunities to pursue HEC related savings, the Company expanded delivery and accessed available additional funds in line with provisions set out in the Guidelines. Spending for the portfolio in 2015 is summarized in Section 10.2 of this report.

➢ Due to the success of HEC in the first half of 2015, budget limitations became more challenging. As a result, Enbridge worked to optimize the DSM budget to accommodate an expansion of the program (relative to the budget), without unduly removing focus from other DSM areas and sectors. With these priorities in mind, and by accessing all of the available options presented to it, the Company determined that the program could not continue to be funded beyond mid-year. As a result, the Company communicated that eligible projects would need to have pre-audits completed in June, 2015 and post audits completed by July 31st, 2015.



- The HEC offer is well-aligned with the Board's guiding principles and key priorities as outlined in the Framework. The offer seeks to reach an increased number of participants, treat customers' homes in a holistic manner, and drive deep savings. In preparation for future growth, the Company assessed the administration for the offer to identify opportunities for improvements. Enbridge determined that a live registration site (to be developed early in 2016) for pre-audits would assist with the management of workflow processes for home audit tracking. With increased participation in the offer anticipated, improvements were necessary to ensure effective processes would be maintained with increased capacity.

5.2 Commercial Resource Acquisition

Enbridge serves over 150,000 Commercial sector customers across the Company's franchise territory. These customers span a wide variety of sub-sectors, which include: Multi-Residential Buildings, Commercial Office Buildings, Schools/Universities, Hotels/Motels, Warehouses, Retail Facilities, Food Services, Hospitals/Health-Care and Government/Municipal Facilities.

Offers designed for commercial customers include custom and prescriptive approaches designed to support the installation of energy efficient equipment and the adoption of energy efficient practices. This is accomplished through the provision of energy audits, technical support, education and incentives.

DSM programming available to commercial customers is delivered directly by Enbridge's Energy Solutions Consultants (ESCs) to customers and building owners/ operators and also through supply chain channels and business partners, including HVAC contractors, engineering firms and energy service advisors.

Table 5.4 2015 Commercial Resource Acquisition Results

Resource Acquisition Commercial Sector	2015 Net CCM	# of Projects [1]	# of Units [2]	TRC-Plus Ratio	PAC Ratio
Custom	350,622,209	562	---	3.25	12.35
Prescriptive	97,416,428	---	16,877	5.71	21.34
Run It Right	2,684,105	28	---	0.33	0.36
Total/Average	**450,722,741**	**590**	**16,877**	**3.39**	**10.78**

1. # of Projects summarizes the number of unique projects for custom offers and RIR.
2. # of Units summarizes the number of units installed for prescriptive offers.

Commercial – Custom and Prescriptive Fixed Incentive Offers

Objectives	The goal of the Commercial Custom offer is to reduce natural gas use through the capture of energy efficiency opportunities in commercial buildings, including retrofits of building components and upgrades at the time of replacement. The offer aims to promote the highest level of energy efficiency. The Commercial Prescriptive offer is designed to capture energy savings in the Commercial sector associated with the installation of prescriptive and quasi-prescriptive technologies.
Target Customer	Both the Custom and Prescriptive offers target commercial customers who are primarily in Rate 6 as well as commercial customers in Rates 135, 145, 110, 115 and 170.
Metrics	As part of the RA program, the primary metric for the Commercial Custom and the Prescriptive offer is lifetime natural gas savings - cumulative cubic meters (CCM) savings.
Tracking Methodology	Savings for each custom project are calculated on an individual basis and results are tracked weekly by the Tracking and Reporting team, utilizing Enbridge's sales tracking software. Data is compiled for Prescriptive offer participants and results are also tracked on a weekly basis by the Tracking and

	Reporting team. All supporting documentation is reviewed for accuracy and completeness and is retained by Tracking and Reporting.
Offer Description	The Custom Commercial offer provides incentives for customers undertaking capital and operational improvements. Typical measures include the installation of high efficiency boilers, controls and building automation systems, heat recovery projects and building envelope improvements. The offer is promoted and delivered by ESCs who are active in the marketplace. ESCs are trusted energy advisors; their technical and energy efficiency sales experience is fundamental to the successful execution of custom projects. Enbridge executes on multiple approaches to reach commercial customers. ESCs work directly with customers, meeting with building operators and facility managers to conduct site visits and educate customers on potential options to improve the energy use of their facilities. They review prescriptive offerings to enable potential upgrade options or present custom recommendations where applicable, based on a building's unique systems and to suit the customer's energy efficiency goals, budgetary considerations and business needs. ESCs also work with national chain and large property management firms, centralizing efforts to introduce savings strategies and align DSM offers with customers' company-wide energy plans. ESCs use their technical expertise to work with smaller firms and managers of standalone buildings by educating them on savings concepts and providing recommendations and savings estimations for potential projects.

	Further, the Company works with a network of business partners to extend outreach to customers and promote awareness of the offers and encourage efforts towards energy efficiency. The Company maintains relationships with service providers (e.g. HVAC contractors, engineering consultants or energy service companies), manufacturers and distributors, ensuring they are well versed about offers and can present savings opportunity scenarios and discuss incentives and application processes with customers. The Commercial Prescriptive offer for 2015 included fixed incentives for various prescriptive and quasi-prescriptive energy efficiency measures impacting space heating, water heating and food service equipment. Prescriptive measures have pre-determined fixed savings based on the size and classification of the equipment. Quasi-Prescriptive measures involve energy savings calculations based on partially pre-determined values, but where one or more variables need to be input in order to determine gas savings for a particular installation. Enbridge offered a full range of prescriptive and quasi-prescriptive measures including:[14] • Demand Control Ventilation (DCV); • Condensing Boilers <300MBH; • High Efficiency Boilers (specified parameters); • Air Doors; • Energy Recovery Ventilation (ERV); • Heat Recovery Ventilation (HRV); • Infrared Heaters; • Condensing Make-Up Air Units; • Ozone Laundry System;

14 Specific details regarding measures included can be found at enbridgegas.com/commercial

	• Low-Flow Showerheads; • Demand Control Kitchen Ventilation System (DCKV); • Energy Star Qualified Dishwashers; • Energy Star Qualified Natural Gas Convection Ovens; • Energy Star Qualified Natural Gas Fryers; • Energy Star steam cookers; and • High efficiency under-fired broilers.
Cost-Effectiveness	Both the Commercial Custom and Prescriptive offers were cost-effective, as supported by the TRC-Plus screening summarized in Table 5.4.
Evaluation Activities	In the case of custom projects, savings for each project are determined with project-specific savings calculations. Where applicable, ESCs utilize standardized engineering calculators developed by Enbridge's technical engineering team. Projects are screened for an additional internal technical review to verify savings calculations as appropriate. Where required, savings calculations are specialized based on project-specific engineering analysis. An independent third-party engineering review, the Custom Project Savings Verification (CPSV), is conducted annually. This verification study has historically consisted of a detailed review of the savings calculations for a statistically representative sample of commercial custom projects.[15] Beginning in 2015, as outlined in the August 21st, 2015 memo from the board (EB-2015-0245), which outlines the new governance structure detailing the OEB's process to evaluate the results of Natural Gas Demand Side Management (DSM) programs from 2015 to 2020, the Board will be responsible for retaining an Evaluation Contractor (EC). The detailed annual evaluation and audit process will be developed as part of the EM&V plan which the EC is expected to draft. The EAC will

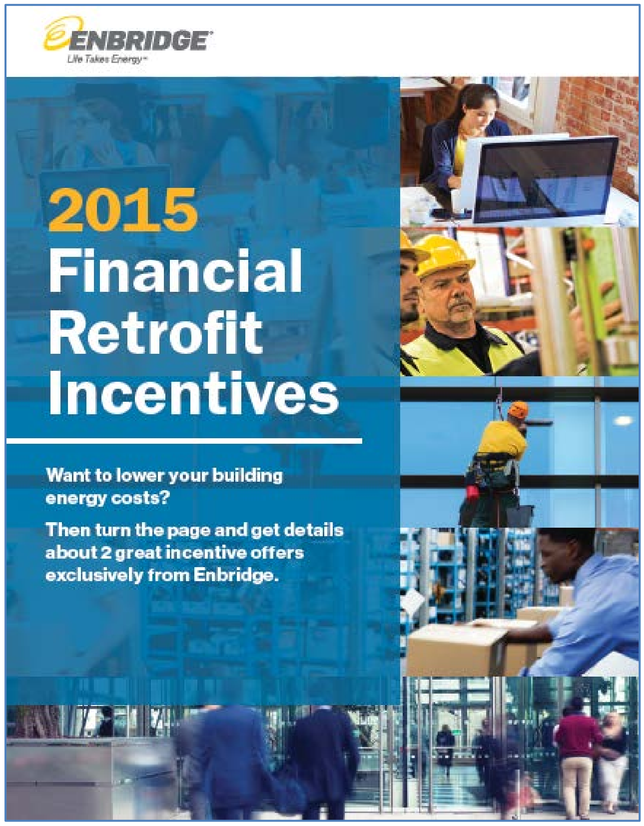
15 The prescribed sampling methodology was developed for Enbridge and Union Gas by Navigant Consulting in 2012, revised in 2014 and endorsed by the TEC. "A Sampling Methodology for Custom C&I Programs", Dan Violette & Brad Rogers, Navigant Consulting, Inc., November 12, 2012. Revised: October 28, 2014

	provide advice and input on the development of the plan as appropriate.
2015 Results	As summarized in Table 5.4, 562 commercial custom projects were claimed in 2015; these projects accounted for 350.6 million CCM in natural gas savings. Custom projects traditionally drive the highest percentage of Commercial results. As per Table 5.4, Commercial Prescriptive measures totaling 16,877 units contributed 97.4 million CCM.

2015 Commentary and Lessons Learned:

- Enbridge continues to provide technical expertise to support and influence Commercial customers and their suppliers to identify and implement capital and operational improvements. Despite challenging rollover targets, natural gas savings results from Commercial DSM efforts were good in 2015.

- With 2015 being a rollover year from the previous multi-year plan, incentives for custom projects remained consistent at $0.10/m^3 of gas saved and fixed incentives specific to prescriptive measures continued both to customers and to contractors/distributors.



- The strongest contributors to commercial custom project results were the Multi-Residential sector, Education and Health-Care sectors.

- Similar to 2014, measures that were among the major drivers to the Commercial prescriptive results in 2015 included prescriptive high-efficiency boilers, infrared heaters, demand control ventilation and ozone laundry systems.
- Competing priorities for Commercial customers continued to be one of the challenges to DSM project uptake in 2015. With limited capital to invest into energy efficiency upgrades, customers must consider a variety of options. For example investing in gas utility DSM initiatives, to decrease their natural gas consumption, versus investing in CDM initiatives, to reduce higher cost electricity consumption. An added challenge for DSM is that customers often stand to benefit from a relatively larger incentive to pursue CDM upgrades on a per energy unit basis.

Upgrading your building by investing in efficient equipment and systems can provide long-term environmental and financial benefits. At Enbridge we know buildings have unique energy requirements and that's why we're providing two great options for you to choose from:

OPTION 1
UPGRADE / REPAIR CURRENT EQUIPMENT

Then look no further than the **Enbridge Custom Retrofit Incentive program**, this option will help you offset the cost of equipment upgrades, major repairs or undertake long overdue maintenance. **Qualified projects receive $0.10 per m³ of natural gas saved (up to $100,000) for implementing energy saving measures.**

QUALIFYING COMMERCIAL RETROFIT PROJECTS MAY INCLUDE:

- Upgrading to high efficiency and condensing boilers
- Improving boiler controls
- Improving ventilation controls
- Installing demand control ventilation
- Installing energy or heat recovery systems
- Installing energy management systems and other automated controls
- Steam trap, economizer or condensate leak repairs
- Insulation upgrades
- Installing reflective panels on radiators
- High extraction washers

OPTION 2
INSTALL NEW EQUIPMENT

Our **Fixed Retrofit Incentive program** is the quick and easy way to offset the cost of installing energy efficiency equipment in your existing buildings.

Exclusive customer incentives for existing buildings

SPACE HEATING	DESCRIPTION	CUSTOMER INCENTIVE AMOUNT
DEMAND CONTROL VENTILATION[1]	Single Zone Retail	$0.04/ft²
	Single Zone Offices	$0.04/ft²
CONDENSING BOILERS[2]	Up to 299 MBH, AFUE of 90% or Greater	$400
HIGH EFFICIENCY BOILERS[3]	Up to 599 MBH	$400
	600 to 999 MBH	$550
	1,000 to 1,500 MBH	$1,400
	1,501 to 2,000 MBH	$2,200
HIGH EFFICIENCY BOILERS FOR SCHOOL BOARDS[4]	Elementary Schools	$1,000
	Secondary Schools	$4,300
AIR DOORS [4,5]	Single Door (Pedestrian Entrance)	$50
	Double Door (Pedestrian Entrance)	$100
	Shipping Doors (8' x 8')	$450
	Shipping Doors (8' x 10')	$600
	Shipping Doors (10' x 10')	$2,400

*Applicable to industrial customers for non-process, space heating applications.

SPACE HEATING	DESCRIPTION	CUSTOMER INCENTIVE AMOUNT
ENERGY RECOVERY VENTILATION [4,6]	Offices, Warehouses & Schools, Minimum Heat Recovery Effectiveness 67%	$0.25/CFM
	Hotels, Restaurants & Retail, Minimum Heat Recovery Effectiveness 67%	$0.40/CFM
	Multi-family, Healthcare & Nursing Homes, Minimum Heat Recovery Effectiveness 67%	$0.75/CFM
HEAT RECOVERY VENTILATION [4,7]	Offices, Warehouses & Schools, Minimum Heat Recovery Effectiveness 61%	$0.20/CFM
	Hotels, Restaurants & Retail, Minimum Heat Recovery Effectiveness 61%	$0.30/CFM
	Multi-family, Healthcare & Nursing Homes, Minimum Heat Recovery Effectiveness 61%	$0.55/CFM
INFRARED HEATERS[4,8]	Single stage up to 300,000 BTUs	$100
	Two stage up to 300,000 BTUs	$200
CONDENSING MAKE-UP AIR UNITS[4,9]	Single speed up to 14,000 CFM, Multi-family & Long-Term Care	Greater of $500 or $0.25/CFM
	Two speed up to 14,000 CFM, Multi-family & Long-Term Care	Greater of $500 or $0.30/CFM
	Single speed up to 14,000 CFM, Other Sectors	Greater of $500 or $0.75/CFM
	Two speed up to 14,000 CFM, Other Sectors	Greater of $500 or $0.25/CFM
WATER HEATING		
OZONE LAUNDRY SYSTEM[10]	Commercial Laundry Washing	$0.006 (x) total annual lbs. of laundry processed
SHOWERHEADS[11]	Multi-family residential buildings only. Replacement of an old showerhead 2.5 GPM or more with a new one at a flow rate of 1.5 GPM or less	$6 per showerhead
FOOD SERVICE		
DEMAND CONTROL KITCHEN VENTILATION SYSTEM	Up to 4,999 CFM, Approved Manufacturers Only	$500
	5,000 to 9,999 CFM, Approved Manufacturers Only	$1,500
	10,000 to 15,000 CFM, Approved Manufacturers Only	$2,500
ENERGY STAR® QUALIFIED DISHWASHERS[12]	Multiple Tank Conveyor, ENERGY STAR® Qualified	$400
	Single Tank Conveyor, ENERGY STAR® Qualified	$250
	Under Counter, High Temperature, ENERGY STAR® Qualified	$100
	Stationary Rack, Low and High Temperature, ENERGY STAR® Qualified	$100
ENERGY STAR® QUALIFIED NATURAL GAS CONVECTION OVENS - FULL SIZE	ENERGY STAR® Qualified	$100
ENERGY STAR® QUALIFIED NATURAL GAS FRYERS	ENERGY STAR® Qualified	$100

If you're interested in improving your existing building's energy performance choose the option that best meets your needs and then **contact your Enbridge Energy Solutions Consultant (ESC)** who will help you through the process.

- Enbridge continued to engage industry stakeholders and organizations in efforts to further support education and build awareness of the Enbridge services and DSM support available. These groups included:
  - The Building Owners and Managers Association (BOMA Toronto, BOMA Ottawa)
  - Restaurants Canada
  - Ontario Restaurant Hotel & Motel Association (ORHMA)
  - Retail Council of Canada
  - Ontario Refrigeration and Air Conditioning (ORAC)
  - The Heating, Refrigeration and Air Conditioning Institute (HRAI)
  - Hotel Engineering/Facilities Manager's Association of Toronto (HEAT)
  - Eastern Ontario Landlord Organizations (EOLO)
  - Association of Condominium Managers of Ontario (ACMO)
  - Canadian Condominium Institute (CCI)
  - Federation of Rental Providers of Ontario (FRPO)
  - Greater Toronto Apartment Association (GTAA)
  - Canadian Healthcare Engineering Society (CHES)

- Ontario Long-Term Care Association (OLTCA)
- Professional Retail Store Maintenance Association (PRSM)
- American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)
- Energy Solution Centre (ESC)
- Continental Automated Buildings Association (CABA)

ENBRIDGE

Case Study

Energy efficient technologies serve up menu of benefits – Boston Pizza

Boston Pizza franchisee Rob Phillips used this Niagara Falls location to test three new energy efficient technologies – a Melink demand control kitchen ventilation system; Enershield air door; and low volume, high intensity kitchen spray valve. The impressive results – in energy cost savings, comfort, and green performance – convinced him to repeat the success at other franchise locations. Enbridge Gas Distribution helped Rob every step of the way with incentives and other assistance.

Challenges

The flagship of Rob Phillips' Boston Pizza franchises is a Morrison Street location in Niagara Falls. It is an open and expansive design with 7,287 square feet of space and high ceilings. As with any restaurant, top-notch heating, ventilation and air conditioning

Rob's original kitchen ventilation system included three conventional range hoods that operated at 100% continuously whether stoves were in full use or not. This meant unnecessary loss of heated make-up air – and higher natural gas bills. The kitchen's hot

➢ Enbridge worked to identify appropriate collaboration opportunities in 2015 that could be leveraged to drive natural gas savings for commercial customers and promote energy efficiency broadly. A Performance Based Conservation initiative with Toronto and Region Conservation Authority (TRCA) involved electricity, gas and water utilities working together with the Independent Electricity System Operator (IESO) in an effort to understand and take action on energy savings opportunities in commercial and institutional buildings. Over the course of three years, the pilot will leverage a new, data-driven methodology to help building owners and managers understand their energy use through benchmarking.

➢ In addition, the Company was active in key industry events and conferences to further build DSM program awareness, and to provide customers with opportunities to discuss their challenges directly with an ESC. Some of these events were:

- City of Toronto – Live Green, Toronto Hotel Sustainability Conference
- Canadian Healthcare Engineering Society, Provincial Trade Show & Education
- Ontario Long Term Care Association, Industry Event
- Toronto and Region Conservation Authority (TRCA), Greening Health Care Event
- Canadian Condominium Institute, Ottawa Conference / Tradeshow
- CivicAction, Race to Reduce
- Eastern Ontario Landlord Organization, Spring Networking Event
- Canadian Federation of Independent Grocers, 2015 Grocery Innovations Conference
- Retail Council of Canada, PM Expo
- Operations, Maintenance & Construction of Ontario Association of School Business Officials Annual Tradeshow

➢ Enbridge continues to work with distributors and contractors to promote the Company's energy efficiency offers and encourage these partners, who are well connected in the market, to help to identify opportunities to encourage customers to consider more energy efficient alternatives. For most prescriptive measures, Enbridge provides a nominal fixed incentive to contractors/distributors.

➢ As outlined in its 2015-2020 Multi-Year DSM Plan, the Company has recognized that current approaches have not had the same impact among smaller, harder to reach customers and segments. Beginning in 2016, the OEB approved a revised approach and separate targets for smaller customers (in terms of average annual gas consumption) distinct from large customers. The company is looking at differentiated marketing and delivery approaches to these groups in order to better respond to the needs of each.

➢ In addition to the formation of a dedicated sales team for smaller customers, the Company intends to expand the industrial online client portal to the commercial sector and develop tools and calculators to be available online to support customers and business partners. Also, the Company is planning a webinar series targeted to smaller commercial and industrial customers and intends to expand the industrial newsletter to include small commercial accounts.

Commercial – Run it Right and Energy Compass

Objectives	The goal of Run it Right (RiR) and Energy Compass is to encourage building owners to improve the energy performance of their buildings through operational improvements and benchmarking. These offers promote the awareness / visibility of building consumption patterns through energy monitoring information services (EMIS), low cost/no cost operational improvement measures and energy savings opportunity assessments. Ultimately, these offers aim to lead commercial customers toward data-driven decision-making.
Target Customer	These offers are targeted to commercial customers in Rate 6, 110, 115, 135, 145 and 170 (with most commercial customers falling in the Rate 6 category). More specifically, the offers are designed for energy managers and building operators of commercial, multi-family and institutional buildings where daily consumption data is accessible. The Energy Compass initiative is marketed to commercial customers that have a portfolio of buildings.
Metrics	As part of the Resource Acquisition program, the primary metric for RiR is lifetime natural gas savings - cumulative cubic meters (CCM) savings. The Energy Compass initiative does not have a scorecard metric.
Tracking Methodology	The 2015 results are based on participants that registered for the RiR offer and completed the implementation of the agreed-upon low/no cost operational measures in 2014. Tracking and Report compiles data for each participant. Applicant information includes site address and building details, also consumption information and meter type are tracked. In addition, details regarding the investigation agent

	conducting the assessment, milestone dates, measures tracked and incentive amounts are recorded. Final regression analysis reports for each participant are maintained and calculated savings are tracked.
Offer Description	The RiR offer, as well as the Energy Compass initiative, is designed to motivate commercial customers towards performance-based conservation. The provision and analysis of detailed energy data aims to allow building operators and managers to make strategic data-driven decisions regarding energy savings and capital investments. Through Energy Compass and RiR, the Company helps commercial customers better manage their buildings, implement operational improvements to achieve energy savings and identify future cost-effective capital improvements. Savings that result from operational improvements implemented in any given year are recorded in the next year, following monitoring and verification. Customers interested in participating in the offer, and meeting the participation criteria, are first engaged by an Enbridge designated investigation agent. This agent conducts a high level energy audit on the participant's facility, identifying a list of operational improvement measures for the customer to implement. Once a customer implements the recommended measures, depending on the complexity of the building systems and annual consumption, a customer is then provided an incentive. Customers are then added to the Enbridge selected EMIS system in order to begin their 12 month monitoring period. Following the 12 month monitoring period, Enbridge provides the customer with a report which summarizes savings.
Cost-Effectiveness	The RiR offer is not cost-effective in 2015, as illustrated by the TRC-Plus and PAC screening summarized in Table 5.4.

	However, the Commercial sector offers overall and the Resource Acquisition program as a whole shows screening results that are cost-effective.
Evaluation Activities	The 2015 results are based on participants that registered for the RiR offer and completed the implementation of the agreed-upon low/no cost operational measures in 2014. A third party firm was retained by Enbridge to determine the 2015 claimed RiR savings. For these participants, gas consumption data for the 12 months prior to implementation (the base year) was used as the base case. Gas consumption was then monitored for 12 months following implementation (the reference year). Gas savings results are based on a standardized statistical regression analysis of actual consumption data for each participant, comparing natural gas consumption during a baseline and reference period. The baseline period is the time period prior to implementation of operational improvements while the reference is the period after improvements. Weather normalization of the baseline and reference gas consumption data is completed.
2015 Results	Results for RiR are based on the calculation of total savings determined in 2015 for participants that enrolled in RiR in 2014. In 2015, volumetric savings of 2.68 million CCM were achieved by the 28 eligible participants. These amounts are outlined in Table 5.4. A further 8 participants were removed from the results due to the inclusion of capital measures during the monitoring period.

2015 Commentary and Lessons Learned:

- As was the case in previous years, an analysis of RiR participant results continues to indicate that average savings levels are significantly lower than the

10% reduction that was suggested in the initial forecast for the 2012 offer design. In 2015, average savings were 4.2%.

- While the number of customers that completed a RiR audit in 2014 was similar to previous years, those that followed through and implemented measures decreased from 53 in 2013 to 36 in 2014. Monitoring and measurement of savings proceeded for these 36 participants. Ultimately gas savings results associated with only 28 participants were included in the 2015 RiR gas savings result. The remaining participants undertook capital projects and therefore were deemed ineligible based on criteria for the 2015 RiR offer.

- In response to the low number of customers who implemented measures relative to the number of customers who showed initial interest and completed an audit, Enbridge sought to increase engagement between the investigation agents and customers signing up for the offer in 2015. Following the customers' receipt of their investigation report, the investigation agent was required to follow up with the customer to provide any assistance to support the implementation of the recommended measures. The expected outcome was that there would be an increase in the number of customers that took action and moved to the monitoring phase. Based on enrollments in 2015, this action has seen an increase in customers proceeding with the monitoring phase.

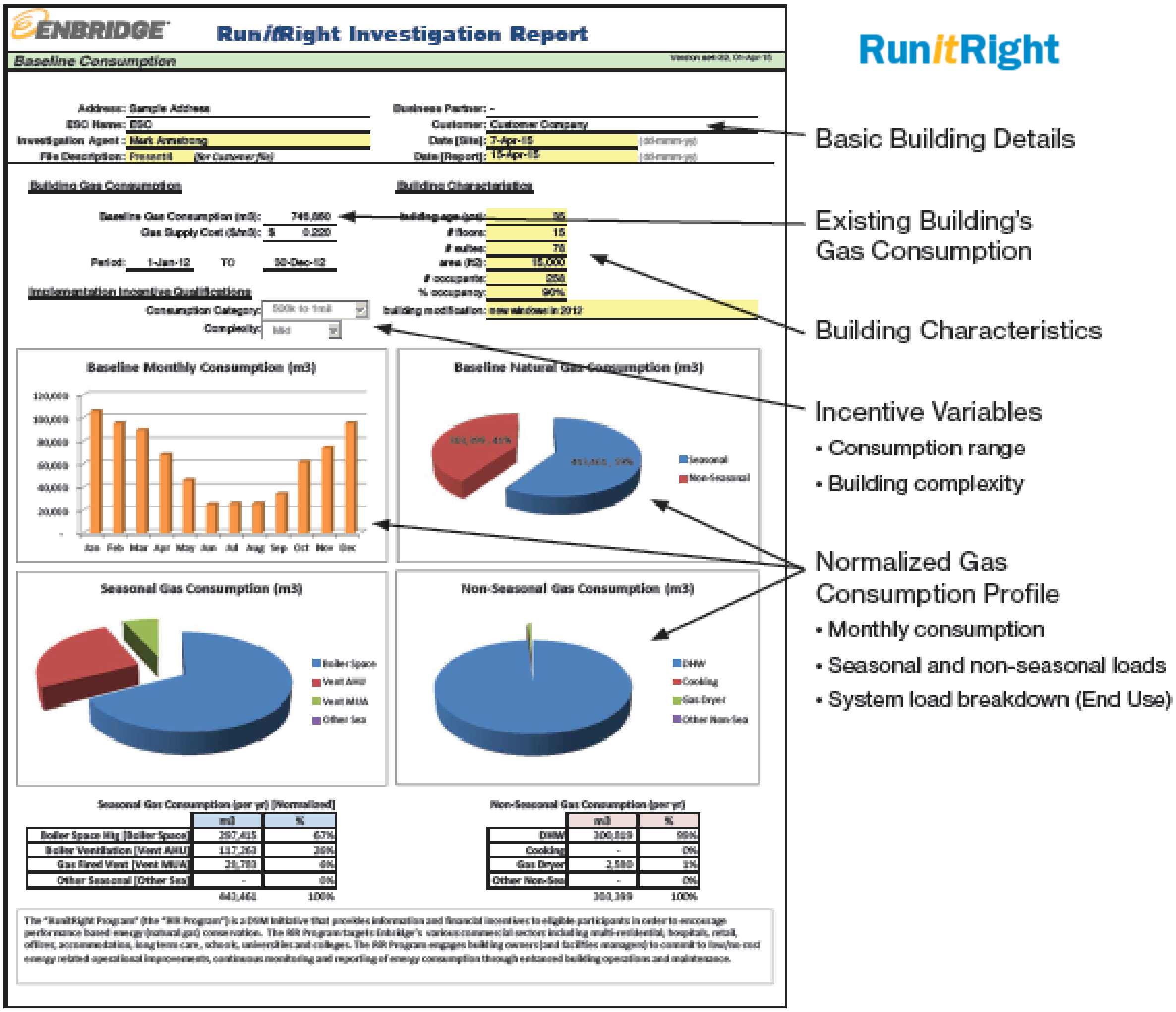
- Enbridge implemented further improvements to support customers enrolled in the offer including:
  - introducing a third-party calling service to educate and generate interest in the offer from existing eligible customers;
  - the development of a Building Automation Systems (BAS) training module ;
  - offering EMIS training to new participants to acclimate them to the software and encourage active usage throughout the 12 month measurement period.
  - The creation of an interactive display to better engage potential customers at various industry events.

- Assessing and interpreting actual results to determine RiR savings remains challenging. Although metered data reflects building consumption, it does not necessarily reflect the building conditions that can change year-over-year and therefore does not always provide a reliable assessment of the savings associated with operational improvements undertaken through the offer.

- An increase or decrease in consumption that occurs as a result of changes in the building not related to operational improvement activities (such as increasing operating hours or building occupancy changes) has an impact on the savings realized through the building's participation in the RiR offer. Such factors can be challenging to monitor and account for in the RiR regression analysis.

- Enbridge has spent considerable time and effort both independently and through the 2013 and 2014 audit processes to explore how to appropriately apply a methodology to capture operational savings. However, the results of these efforts have proved inconclusive to date. In a continued effort to further inform an appropriate methodology to be used, the Company intends to implement quarterly energy logs with participants to better understand building condition changes.

- As specified in the 2015-2020 Multi-Year DSM Plan, beginning in 2016 RiR has been redesigned to allow for the calculation of operational improvements even where the customer intends to proceed with capital projects. This revision should further remove barriers to participation.



5.3 Industrial Resource Acquisition

Industrial – Custom Solutions and Prescriptive Fixed Incentives Offers

Objectives	The Industrial Custom Solutions offer is designed to capture cost-effective energy savings within the Industrial sector by delivering customized energy solutions aimed at supporting customers through a continuous improvement approach. Industrial Energy Solutions Consultants (ESCs) focus on assisting customers with the adoption of energy efficient technologies by overcoming financial, knowledge or technical barriers. The Industrial Prescriptive offer aims to capture energy savings in the Industrial sector by installing applicable prescriptive and quasi-prescriptive technologies, with a focus on increasing the adoption of energy efficient technologies among small industrial customers.
Target Customer	Both the Custom Solutions and Prescriptive offers are available to industrial customers (including Agricultural customers) in Rates 6, 110, 115, 135, 145 and 170. Custom projects encompass opportunities where savings are linked to unique industrial processes, building specifications, uses and technologies. With the Custom Solutions offer, Enbridge primarily targets industrial customers (both large and small) with significant process loads and high annual consumption. The technologies targeted to customers included in the prescriptive offer are often most suitable to smaller industrial customers whose gas usage is less weighted to the high process load profiles typical in larger industrial customers and who proportionally have higher seasonal gas usage.

Metrics	As part of the Resource Acquisition program, the primary metric for the Industrial Custom and the Prescriptive offer is lifetime natural gas savings - cumulative cubic meter (CCM) savings.
Tracking Methodology	Savings for each custom project are calculated on an individual basis and then tracked weekly by the Tracking and Reporting team, utilizing Enbridge's sales tracking software. Data is compiled for Prescriptive offer participants and also tracked on a weekly basis by the Tracking and Reporting team. All supporting documentation is reviewed for accuracy and completeness and is retained by Tracking and Reporting.
Offer Description	In the Industrial sector, the Continuous Energy Improvement (CEI) approach includes the Industrial Custom Solutions offer and the Prescriptive offer together with a number of enabling initiatives, such as support for industrial customers in identifying energy-saving opportunities through to assistance with project implementation. These offers are primarily promoted and delivered by ESCs (professional engineers) who are active in the marketplace. ESCs are trusted energy advisors that work with customers to determine solutions to address multiple objectives, namely production, energy efficiency and budgetary considerations. Work involves addressing technical barriers to energy efficiency adoption as well as financial barriers that may hinder business justification and implementation. Enabling initiatives allow ESCs to work with the customers to identify potential opportunities, quantify benefits, and justify action. Such initiatives include: ESCs leveraging their skills and tools to identify efficiency opportunities; involvement of third-party vendors to conduct specific types of audits or assessments of facilities; and/or ESCs assisting with the development of project implementation plans.

	Due to the unique nature of industrial customers, custom solutions developed by ESCs are designed and engineered to meet the specific requirements of each particular customer's facility. Five core components are common to the Custom offer: **Knowledge Development**: Technical publications, quarterly updates, themed workshops and a resource based energy solutions portal are offered to provide customers with the knowledge to make informed decisions through education. **Opportunity Identification**: ESCs provide support to assist customers in the identification of efficiency opportunities, such as equipment testing and assessment and thermal imaging. **Measurement**: ESCs assist customers in selecting appropriate means of measurement to quantify key energy inputs. **Engineering Analysis**: ESCs assist customers who do not have the resources needed to conduct financial, technical and enterprise risk evaluations for potential projects. **Implementation Support**: ESCs work with customers on an implementation plan and connect them with business partners to complete the project. The following tiered incentive structure which was introduced in 2014 was once again offered in 2015 with the custom offer: $0.20/m^3 for first 50,000 m^3 gas saved $0.05/m^3 for gas savings above 50,000m^3 This incentive structure was designed to provide additional support to customers (both large and small) with the implementation of smaller projects. The Industrial Prescriptive offer evolved by leveraging existing Commercial offers applicable to the industrial customer base. The Industrial Prescriptive offer incorporates a fixed incentive approach and includes incentives designed to help offset the cost of energy efficiency upgrades specifically relevant to industrial facilities such as Air Doors, Heat Recovery

	Ventilators, Energy Recovery Ventilators, Condensing Make-up Air Units, Infrared Heaters and Destratification Fans.
Cost-Effectiveness	Enbridge continues to demonstrate a high level of cost-effectiveness for the Industrial sector offers as supported by the TRC-Plus and PAC screening summarized in Table 5.5 that follows.
Evaluation Activities	In the case of custom projects each project is assessed individually. Subsequent to project-specific savings calculations being completed by ESCs, an internal technical review of project applications and savings calculations is conducted. ESCs utilize standardized engineering calculators developed by Enbridge's technical engineering team. Where required, savings calculations are specialized based on project-specific engineering analysis. An independent third-party engineering review, the Industrial Custom Project Savings Verification (CPSV), is conducted annually. This verification study has historically consisted of a detailed review of the savings calculations for a statistically representative sample of Industrial sector custom projects.[16] Beginning in 2015, as detailed in the August 21st, 2015 memo from the board (EB-2015-0245), which outlines the new governance structure detailing the OEB's process to evaluate the results of Natural Gas Demand Side Management (DSM) programs from 2015 to 2020, the Board will be responsible for retaining an Evaluation Contractor (EC). The detailed annual evaluation and audit process will be developed as part of the EM&V plan which the EC is expected to draft. The EAC will provide advice and input on the development of the plan as appropriate.

16 The prescribed sampling methodology was developed for Enbridge and Union Gas by Navigant Consulting in 2012, revised in 2014 and endorsed by the TEC. "A Sampling Methodology for Custom C&I Programs", Dan Violette & Brad Rogers, Navigant Consulting, Inc., November 12, 2012. Revised: October 28, 2014

2015 Results	There were 115 projects completed in the Industrial custom offer in 2015, and contributing 173.4 million CCM. Custom projects for industrial customers can be varied across a wide range of upgrades and improvements. In 2015, results from custom projects were led by savings from projects focused on industrial process efficiency improvements, the installation of control systems, and improvements to operational processes unique to specific customers. Prescriptive results totalled 7.59 million CCM and included 235 units installed. The focus for the Industrial prescriptive technologies in 2015 was Air Curtains and Infrared Heaters.

Table 5.5 2015 Industrial Resource Acquisition Results

Resource Acquisition Industrial Sector	2015 Net CCM	# of Projects [1]	# of Units [2]	TRC-Plus Ratio	PAC Ratio
Custom	173,397,871	115	---	6.21	15.00
Prescriptive	7,593,008	---	235	5.03	51.04
Total/Average	**180,990,879**	**115**	**235**	**6.15**	**15.45**

1. # of Projects summarizes the number of unique projects for custom offers.
2. # of Units summarizes the number of units installed for prescriptive offers.

2015 Commentary and Lessons Learned:

- The industrial sector utilizes most of its energy for process related consumption as opposed to heating and ventilation purposes. Plants consume a small portion of energy compared with the process equipment within the facility. Many industrial customers lack technical knowledge regarding energy efficient technologies that may help improve these processes and reduce overall energy consumption. Consequently, the industrial team focuses its efforts on helping customers identify ways to improve efficiency with process lines and the optimization of operational procedures.

- Overall, the Custom Solutions offer remained largely unchanged in 2015 from the previous year. Results continued to reflect the developing trend seen over the last few years with a shift from capital-intensive projects such as equipment upgrades, to opportunities focused on process improvements. The outcome has been an increasing proportion of projects which tend to yield good annual savings but lower CCM.

- Though the industrial team has identified an increasing number of potential opportunities year over year, the associated savings generated from completed projects has decreased in terms of cumulative gas saving results.

- Custom projects tend to be resource intensive requiring extensive technical expertise and data analysis; whereas prescriptive, fixed incentive projects are less complex to execute, and therefore a good alternative for smaller customers. The Company continued to leverage a distribution network of business partners and service providers to assist in the promotion of the Prescriptive offer.

- 2015 prescriptive project customers benefitted from financial incentive support for the installation of Infrared Heaters as well as Air Doors. In total, 235 prescriptive projects were completed.

- The Company continues to pursue opportunities to undertake audits and studies at industrial customers' facilities (e.g. plant energy assessments, steam trap audits or meter studies) to identify for the customer potential savings that could be realized with the implementation of various improvements. Approximately 70 audits were completed in 2015.

- Enbridge offered a variety of materials and forums aimed at increasing awareness of energy efficiency opportunities and benefits, educating industrial customers and providing resources to research and evaluate potential improvement solutions. Efforts in 2015 focused on a number of initiatives including:

  - Energy efficiency workshops and webinars;
  - Quarterly newsletters (via email blasts);
  - Audits and Assessments; and,
  - Industrial Energy Solutions Portal

- Over the course of 2015, to further increase awareness of energy efficiency in customers' facilities, the industrial team hosted workshops focused on educating customers and their employees on identifying energy conservation opportunities and providing information to help evaluate potential projects. These workshops helped customers identify projects that not only resulted in natural gas savings, but also identified electric and water savings opportunities. The 2015 workshops included the following:

  - **Process Heating Efficiency Workshop**
    Focused on helping customers understand and identify process heating related energy savings opportunities and discover how incremental changes can generate real savings.

ENBRIDGE

enbridgegas.com/industrial

We are here to help:
Energy Solutions Consultants
1-888-427-8888
energysolutions@enbridge.com

Enbridge offers industrial customers free services and financial incentives designed to help you make the most of the energy you consume.

Process Heat Industrial Workshop

Process heating is vital to nearly all manufacturing processes, supplying the heat needed to produce basic materials and commodities.

Are you optimizing your process heating equipment and its settings? If not, sign up today for our FREE workshop, led by Damir Naden, Energy Efficiency Manager, Enbridge Gas Distribution.

Workshop Details

Date: June 9, 2015
Location: International Centre, 6900 Airport Road, Mississauga
Time: 8:30 – 11:00 a.m. Complimentary breakfast will be served at 8:00 a.m.
Cost: Free for Enbridge Industrial customers

Space is limited!
Click Here to Register

The workshop will help Plant Managers, Energy Efficiency Managers, Facilities or Maintenance Managers identify opportunities to save energy and to learn more about equipment dedicated to process heating in manufacturing

Workshop Topics Include:

- Process heating basics
- Thermal heat transfer basics
- Heat recovery
- Industrial oven optimization
- Hands-on examples how to turn the theory into real life savings

Enbridge Gas Distribution will use e-mail addresses submitted only for the purpose of sending communications regarding its available industrial programs. Subscribers may opt out of receiving Enbridge Gas Distribution's industrial program communications at any time by clicking Unsubscribe.

Privacy Policy: [illegible]
Enbridge Gas Distribution P.O. Box 650, Scarborough, Ontario, M1K 5E3

  - **Successful Energy Management Workshop**
    Provided attendees with the opportunity to learn more about energy management planning and how to develop a solid energy baseline of their facilities.
  - **Heating and Ventilation Workshop**
    This session focused on educating customers on how to recognize the symptoms of negative pressure in their facilities, calculate the associated operating costs, and identify no cost/low cost solutions to improve

efficiencies, save money and create a more comfortable work environment for employees year round.

- **Energy Management Success Stories**
  This workshop featured speakers from two leading Ottawa area customers who shared how energy management helped them find and implement energy efficiency solutions that resulted in significant emission reduction and cost savings.
- **Boiler Efficiency Workshop**
  Enbridge's first workshop held at a customer site. Attendees were provided with a site tour of the host client's facility where they were able to explore practical applications of how to optimize energy efficiency, improve productivity and significantly reduce emissions and operating costs in their boiler plant.

- ➢ Over 120 participants took part in these workshops in 2015 and most workshop participants attended more than one event, which serves as an indication that these customers value the information provided. Workshop feedback survey results were excellent with ratings of 95% support in terms of relevancy of the content covered.

- In an effort to drive efficiency projects, limited time incentive campaigns are now introduced at workshops. For example, as part of the HVAC workshop, a campaign to cover the cost of the purchase and installation of a centralized control system for exhaust fans was introduced. In association with the Boiler Efficiency Workshop, customers were offered double the regular incentive for any boiler related energy efficiency project completed within a limited time period. These campaigns will be continued in 2016.

- The Company has established and developed solid relationships with many of the larger industrial customers; however the Company has recognized there is more work needed to improve engagement and develop contacts with the smaller industrial customer base. In 2015, Enbridge worked in collaboration with EnerSource on a Collaborative Energy Assessment initiative targeting the smaller customer segment in both the commercial and industrial sectors. The Company leveraged a third party vendor to connect with smaller commercial and industrial customers and offer them a free energy assessment. As part of this effort, the third party vendor would collect specific customer contact, business and facility information to inform the development of targeted strategies and offers that would more effectively meet the needs of this underserved customer base. The outreach initiative also served as a means of promoting tools, incentives and offers currently available through Enbridge, to a segment of the sector that was previously not well engaged.

- The Industrial Energy Solutions portal launched in 2014 continued to evolve. The portal provides industrial customers, contractors and business partners with the tools to:
  - Identify and quantify energy efficiency opportunities
  - Calculate energy savings
  - Apply for Enbridge financial incentives
  - Learn about different types of energy efficiency technologies
  - Request support from an Energy Solutions Consultant

- Several enhancements were made to the portal in 2015 including:
  - New calculators for the following technologies:
    - Air compressor heat recovery
    - Condensing economizers

    - Feedwater economizers
    - Low temperature chemical wash
    - Exhaust reduction
  - The addition of recorded workshop videos with online access for customers who were unable to attend the events in person.

- In accordance with the rollover for 2015, as in prior years and as outlined in the DSM plan (EB-2015-0049), budget spending on programs and activities for rate classes 110, 115 and 170 was capped. "The purpose of these limits is to ensure that the maximum cost to be borne by industrial customers in these rate classes is known in advance and capped."[17]

- Table 5.6 details the actual spending (including allocated overheads but excluding Low Income Allocations) relative to prescribed spending limits for each rate class and shows that spending is within the limits set out for all three rate classes.

Table 5.6 Rate Class 110, 115 and 170 Spending Limits vs. 2015 Actual Spending

Rate Class	2015 Spending Limit	2015 Spending*
110	$1,721,000	$1,424,754
115	$1,333,000	$478,971
170	$2,264,000	$295,557

****2015 Actual Spending amounts include program costs and overheads but exclude Low Income LEAP allocations.***

- Both of the industrial custom and prescriptive offers continue to be important components in Enbridge's DSM portfolio and will be continued in 2016. As outlined in the Company's 2015-2020 Multi-Year DSM Plan, the Company continues to pursue strategies to successfully drive savings within the smaller industrial customer group. The Company will continue to look at ways to tailor efforts to realize achievement in this challenging market segment.

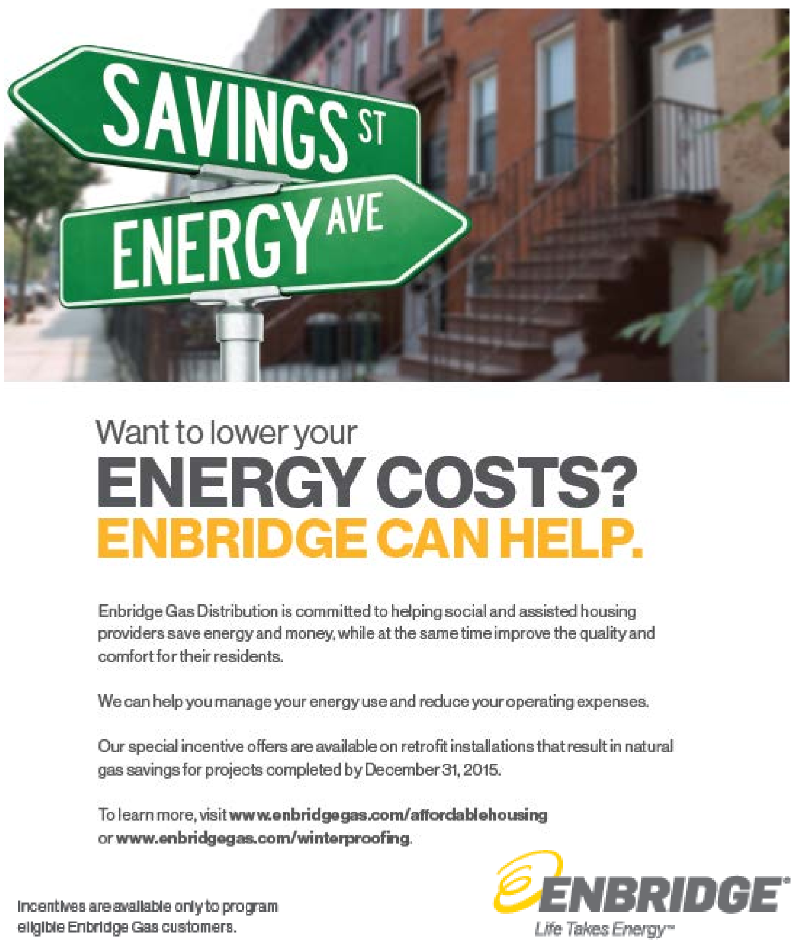
17 Enbridge Gas Distribution Inc. 2015-2020 Multi-Year DSM Plan, OEB File: EB-2015-0049, Exhibit B, Tab 1, Schedule 3, Page 7 of 19.

6. Low Income Scorecard

Enbridge is a leader in the delivery of energy efficiency programs specifically designed for low income customers. Programming has evolved considerably since DSM activities for this market were first offered in the Enbridge franchise in 2004.

Enbridge's Low Income offers are similar to Resource Acquisition offers in that they consist of the installation of energy efficient equipment or measures. However Low Income offers are set apart to recognize the unique needs of their target customer base. Though these offers may result in a lower benefit/cost ratio – Total Resource Cost – than similar offers delivered to non-low income customers, they are designed to address the needs of these consumers and include other important societal benefits.

Performance in terms of the Low Income scorecard for 2015 is measured primarily in terms of net CCM of natural gas savings, however also includes a metric based on program enrollment.



The Low Income program focuses on helping to reduce the energy costs facing low income customers and their housing providers through the installation of measures and thermal envelope improvements to achieve water and space heating savings. Design and delivery considerations for this segment are unique from traditional approaches. As such, approaches are adopted to best reach out to these vulnerable customers and raise customer awareness, encourage resident and building staff engagement, and in turn, build participation. This community includes low wage households, seniors, recent immigrants to Canada and often people with special needs. The Low Income program comprises two segments: Single Family Residential (Part 9) buildings and Multi-Residential (Part 3) buildings.

Enbridge's delivery strategy for the Low Income sector focuses on leveraging available channels and resources, community-based organizations (CBOs) and local community service providers. These groups have established relationships with trusted organizations that support the social service needs (housing affordability and environmental sustainability) of low income consumers.

The Company has also been particularly effective in building collaborative partnerships in the marketplace with LDCs and municipalities. Enbridge has recognized the benefits of collaboration with these partners, as well as with social and assisted housing support networks, in helping to inform and improve program delivery. Proactive stakeholder and customer relationship management provides for continuous program improvement and refocusing of program strategies to be responsive to housing providers' needs and the evolution of affordable housing.

In the social housing space, a key partner in the Enbridge franchise area is Toronto Community Housing (TCH). As the largest social housing provider in Canada and the second largest in North America, TCH provides homes to roughly 60,000 low income households.

The Low Income program exhibited strong results in 2015 relative to scorecard performance targets. Results in the Single Family (Part 9) segment were strong, totaling 28.07 million CCM, surpassing the middle (100%) target. In the Multi-Residential (Part 3) segment, results totalled 63.97 million CCM in natural gas savings.

Table 6.0 2015 Low Income Scorecard

Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Single Family (Part 9)	*Cumulative Savings (million m^3)*	***50%***	18.10	24.10	30.20	**28.07**
Multi-residential (Part 3)	*Cumulative Savings (million m^3)*	***45%***	51.60	68.70	86.00	**63.97**
Multi-residential (Part 3) LIBPM [1]	*Percent of Part 3 Participants Enrolled* [2]	***5%***	30%	40%	50%	**65%**

1. LIBPM - Low Income Building Performance Management is the Low Income offer complement to the Commercial Run It Right (RIR) offer.
2. Low Income Building Performance Management (LIBPM) percentage of Part 3 buildings enrolled in current year program = (x+y)/(x+y+z):
x = # of new LIBPM buildings in the current year that have participated in another aspect of the Low Income program in a previous year of 2012-2014 plan; y = # of new LIBPM buildings in the current year that have not previously participated in the Low Income program; z = # of buildings in the current year that have implemented custom projects other than LIBPM.

Table 6.1 2015 Low Income Results

Low Income Component	CCM 100% Target (m^3)	2015 Net CCM (m3)	TRC-Plus Ratio	PAC Ratio	# of Projects [1]	# of Units [2]
Single Family (Part 9)	24,100,000	28,067,264	1.06	0.96	1,343	1,102
Multi-Residential (Part 3)	68,700,000	63,969,353	3.20	4.76	96	3,331
Total/Average	**92,800,000**	**92,036,617**	**1.88**	**2.00**	**1,439**	**4,433**

1. # of Projects summarizes the number of unique projects Home Winterproofing and for custom offers.
2. # of Units summarizes the number of units installed for prescriptive offers.

All Low Income offers delivered to Enbridge customers in 2015, with the exception of the Low Income Buidling Performance Management offer, will be continued in the Low Income DSM program in 2016. Details regarding individual offers are discussed below.

6.1 Single Family (Part 9)

Home Winterproofing and Prescriptive Measures

Objectives	The goal of the Single Family Low Income offer is to enable energy savings through the reduction of hot water use and space heating demand in low income single family households through the installation of thermal envelope improvements, space heating and water saving measures.
Target Customer	This offer targets Rate 1 homeowners and tenants living in low-rise homes within the Enbridge franchise area who need assistance with their energy costs. Income verification is a requirement for participation in this offer. Eligible customers must meet the following criteria: • Income is at or below 135% of Statistics Canada's Low Income Cut-Off (LICO); • Occupants of single detached and low-rise multi-family (3 stories or less); • Private homeowner or tenant who pays their own gas bills; or • Tenants residing in social and assisted housing, regardless of gas bill payment responsibility.
Metrics	The primary metric is cumulative cubic meter (CCM) savings.
Tracking Methodology	In the case of Home Winterproofing, reports are submitted from delivery agents summarizing installation site information (e.g., address, ownership, housing type) and natural gas savings (m^3) calculated based on the results of customized energy audits conducted by energy auditors for income qualified participants.

	Participation also is tracked by type of tenancy (i.e., social housing or privately-owned dwellings). Similarly, monthly reporting is provided by delivery agents and summarizes unit installations for each prescriptive measure installed. Monthly reports are compiled by the Tracking and Reporting team, utilizing Enbridge's sales tracking software.
Offer Description	The Low Income Home Winterproofing offer is available for: • qualified Part 9 buildings (three stories or less); • private homeowners and residential tenants within the Enbridge franchise who meet the established income eligibility criteria; • residents of social housing; and • recipients of social assistance benefits. For each Part 9 single family home, Enbridge aims to comprehensively address all cost-effective opportunities, provided that the customer accepts all such measures. Basic prescriptive measures including showerheads, aerators, programmable thermostats and heat reflector panels are offered. The Winterproofing offer provides low income customers with a free home energy audit and upgrades that may include: attic, wall and/or basement insulation, door and window caulking and draft-proofing. Enbridge's main approach to delivering the Winterproofing offer is to work with experienced and reliable delivery agents who perform the energy audits and install measures. Upgrades are determined by a free home energy audit performed by a Certified Energy Auditor to determine which cost-effective measures are most appropriate for each home. Basic measures, as defined above, are offered as part of the screening process. Once the measures are installed, a second

	home energy audit is conducted to calculate the gas savings realized. EnviroCentre, Green Communities, and GreenSaver continued as the three primary service providers contracted by Enbridge to market and deliver the offer. These delivery agents are well established in their communities with recognized connections to low income proponents throughout the franchise area. The strategy of delivering the offer in partnership with community-based organizations with strong links to social service agencies has proven to be an effective way of connecting with a hard-to-reach customer segment. Where possible, delivery agents also refer participants to the local electric utility's conservation weatherization program.
Cost-Effectiveness	Low Income programs are often amongst the most expensive to deliver. As per the Guidelines, the Low Income program screening threshold is 0.70; the Low Income Part 9 offer was cost-effective as supported by the TRC-Plus and PAC screening in Table 6.2.
2015 Results	Single Family (Part 9) results were solid in 2015. Actual cumulative savings were 28.07 million CCM, as outlined in Table 6.2. These results exceeded the middle (100%) target of 24.1 million CCM set out in the 2015 DSM Plan. The Enbridge Home Winterproofing offer reached 1,343 low income households in 2015. Many of these homes also received basic prescriptive measures including showerheads and aerators where appropriate, and in some cases also benefitted from the installation of heat reflector panels.

Table 6.2 2015 Single Family (Part 9) Low Income Results

Low Income Component	2015 Net CCM (m3)	TRC-Plus Ratio	PAC Ratio	# of Projects [1]	# of Units [2]
Single Family (Part 9)	28,067,264	1.06	0.96	1,343	1,102

1. # of Projects summarizes the number of unique projects for Home Winterproofing.
2. # of Units summarizes the number of units installed for prescriptive offers.

2015 Commentary and Lessons Learned:

- With Green Light on a Better Environment (GLOBE) no longer a delivery agent due to internal structuring at Housing Services Corporation (HSC), Enbridge reallocated the social housing customers to the remaining delivery agents to expand their customer coverage. Enbridge continues to work with HSC as an energy champion within the sector.

- The combined efforts of delivery agents servicing the privately-owned low income housing market, coupled with continuing work done in social housing (Part 9) properties culminated in a strong result in 2015 with 1,343 homes benefitting from the Home Winterproofing offer in 2015.

- The Company is particularly pleased with the results that were accomplished in 2015, through its work with Ottawa Community Housing (OCH), the second largest housing provider in Ontario. Enbridge worked diligently in managing the performance of EnviroCentre – the Ottawa area delivery agent – to ensure that Enbridge was responsive to the needs of OCH and their residents' while at the same time achieving savings targets.

- As summarized in Table 6.3, 56% or 757 projects claimed in 2015 involved privately-owned houses, the remaining 586 or 44% of homes were social housing. On average, CCM savings per home averaged 20,795 CCM for both social housing buildings and privately-held dwellings.

Table 6.3 Home Winterproofing – Breakdown of Results

2015 Home Winterproofing Results		
# Social Housing Projects	586	44%
# Privately-Owned Projects	757	56%
Total Number of Projects	**1,343**	**100%**
Average CCM Savings/Home - Social Housing (m3)	21,066	
Average CCM Savings/Home - Privately-Owned (m3)	20,585	
Average CCM Savings/Home - All Projects (m3)	**20,795**	

- Toronto Hydro, in co-operation with Enbridge, submitted a business case to IESO to develop a pilot program for joint delivery of the gas and electric Single Family Low Income programs. The pilot is intended to identify cost-efficiency opportunities for joint delivery and enhancements in customer experience. The intention is that this effort will provide a blueprint for a jointly delivered province-wide program.

- Significant efforts in 2015 focused on collaborating with Toronto Hydro to develop a joint initiative between the two utilities to deliver their respective Low Income Single family offers – utilizing one common delivery channel within the City of Toronto.

- Through the Home Assistance Program (HAP) sub-committee of the IESO Residential Working Group, Enbridge worked with IESO in the development of the business case for an updated HAP program in 2015. Of note, these efforts resulted in a streamlined application process, with one single application required for both gas and electric programs. In addition, the HAP income qualification approach was revised so that participants who had already qualified for the gas offer were automatically eligible for HAP. The new HAP program was also revised to incorporate a similar pricing approach used by Enbridge for the implementation of the audit and measures.

- The successful delivery of Home Winterproofing to Toronto Community Housing (TCH) required that efforts integrate with TCH's overall building repair and energy efficiency action plans. A thorough assessment within various TCH departments and Enbridge spanned several months and resulted in a signed Memorandum of Understanding between TCH and Enbridge. A steering committee including Enbridge, TCH, Toronto Hydro and GreenSaver was created with the objective of working to prioritize the delivery of the electricity and gas programs for 2016 and beyond.

➢ Promotion of the Home Winterproofing offer through webinars and information sessions facilitated by the delivery agents to social agencies and community groups continued in 2015. Specific, marketing and sponsorship efforts included:

- buck slips (including a jointly produced piece by Enbridge and Toronto Hydro) and postcards for delivery agents to use at community centre events, social agencies and direct mail across the franchise area;

- street posters were posted outside variety stores in identified low income communities to help increase participation;
- a collaboration effort with the Canadian Health Media Network placed Home Winterproofing brochures and videos in 146 medical offices (resulted in over 800,000 impressions);
- expansion of social media efforts, including digital advertising with Metroland and Google was new in promoting awareness across various channels; and,

- sponsorship through LIEN, HSC and ONPHA to promote initiatives to social and affordable housing providers including:
  - LIEN annual conference
  - ONPHA annual conference inclusive of tradeshow booths and workshops
  - ONPHA Regional Meetings
  - HSC Energy Forum
  - Sponsored Stories
  - E-Alert Advertising

➢ Enbridge continued to engage in training and quality control efforts with delivery agents to ensure good work plan documentation and submission requirements were maintained to support tracking and reporting. Enbridge also facilitated focus groups in Toronto, Niagara, and Peterborough with local program participants. This provided an opportunity for past participants to provide feedback and recommendations for: customer experience improvements; channels for communication; and, marketing messages that led to participation.

➢ The LEAP outbound calling campaign continued for 2015. An estimated 10% of LEAP participants that Enbridge attempted to contact were ultimately transferred to a delivery agent in their area to discuss the Home Winterproofing opportunity. Enbridge continues to talk to LEAP agencies with the objective of allowing Enbridge to engage immediately with participants at the time of LEAP application for inclusion in Home Winterproofing.

➢ The Low Income Home Winterproofing offer will continue to be an important focus for Enbridge in 2016.

6.2 Multi-Residential (Part 3)

Custom Projects and Prescriptive Measures

Objectives	The goal of the Multi-Residential Low Income offer is to enable energy savings through the reduction of space heating demand and hot water use in low income multi-residential buildings through the installation of thermal envelope improvements, space heating and water saving measures.

Target Customer	This offer targets multi-residential social housing providers and managers. The offer also targets eligible owners and property managers of privately-owned multi-unit residential buildings (MURBs) in the City of Toronto, which provide housing to a market that includes low income customers and families based on screening criteria established in collaboration with Enbridge's Low Income Consultative Working Group.
Metrics	The primary metric is cumulative cubic meter (CCM) savings.
Tracking Methodology	As with Commercial custom projects, the savings for each custom project are calculated on an individual basis. Additionally, savings per unit installed for each type of prescriptive measure are tracked and totalled. Results are recorded and summarized through a monthly tracking process utilizing Enbridge's sales tracking software. All supporting documentation is reviewed for accuracy and completeness and is retained by Tracking and Reporting.
Offer Description	Low Income Multi-Residential (Part 3) efforts help social housing providers and MURB managers improve the energy efficiency of aging buildings. The Low Income Multi-Residential offer takes a "building as a system approach" to energy efficiency. It targets housing providers, building operators and tenants with a range of measures such as equipment replacement, thermal envelope improvements and controls, and includes enhanced financial incentives, technical information services, building assessments/audits, education and project facilitation.

	Financial barriers inherent in the Low Income sector related to limited capital availability are addressed by providing an increased financial incentive relative to the standard custom offer; incentives are based on annual natural gas savings up to $100,000 or 50% of project cost. Prescriptive equipment replacement is incented at a set dollar amount depending on efficiency levels. These measures include specific condensing/high efficiency boilers, energy recovery ventilation systems and heat recovery ventilation systems. A free direct install showerhead installation program is also available. Technical issues are addressed by engaging sector experts to provide a suite of services including benchmarking, energy audits, technical assistance and project facilitation. Financial subsidy is provided towards energy audits, building and equipment inventories, and consumption monitoring activities. Direct install in-suite measures, low-flow showerheads and heat reflector panels are provided for eligible buildings.
Cost-Effectiveness	As per the Guidelines, the Low Income program TRC-Plus screening threshold is 0.70. The Low Income Part 3 offer was cost-effective as supported by the TRC-Plus and PAC screening – see Table 6.4.
Evaluation Activities	In the case of custom projects, savings for each project are determined with project-specific savings calculations. Where applicable, ESCs utilize standardized engineering calculators developed by Enbridge's technical engineering team. Projects are screened for an additional internal technical review to verify savings calculations as appropriate. Where required, savings calculations are specialized based on project-specific engineering analysis.

	An independent third-party engineering review, the Custom Project Savings Verification (CPSV), is conducted annually. This verification study has historically consisted of a detailed review of the savings calculations for a statistically representative sample of Commercial/Low Income custom projects.[18] Beginning in 2015, as outlined in the August 21st, 2015 memo from the board (EB-2015-0245), which outlines the new governance structure detailing the OEB's process to evaluate the results of Natural Gas Demand Side Management (DSM) programs from 2015 to 2020, the Board will be responsible for retaining an Evaluation Contractor (EC). The detailed annual evaluation and audit process will be developed as part of the EM&V plan which the EC is expected to draft. The EAC will provide advice and input on the development of the plan as appropriate.
2015 Results	The Low Income Part 3 Multi-Residential offer achieved 63.97 million CCM natural gas savings in 2015.

Table 6.4 2015 Multi-Residential (Part 3) Low Income Results

Low Income Component	2015 Net CCM (m3)	TRC-Plus Ratio	PAC Ratio	# of Projects [1]	# of Units [2]
Multi-Residential (Part 3)	63,969,353	3.20	4.76	96	3,331

1. # of Projects summarizes the number of unique projects for custom offers.
2. # of Units summarizes the number of units installed for prescriptive offers.

2015 Commentary and Lessons Learned:

- The Low Income sector faces inherent financial barriers due to limited capital availability, therefore an increased financial incentive relative to the standard custom offer is provided. Projects in the Low Income sector are generally incented based on \$0.40/m^3 of gas saved for custom measures including building envelope, fans, boilers, heat recovery/economizers and make-up air units.

18 The prescribed sampling methodology was developed for Enbridge and Union Gas by Navigant Consulting in 2012, revised in 2014 and endorsed by the TEC. "A Sampling Methodology for Custom C&I Programs", Dan Violette & Brad Rogers, Navigant Consulting, Inc., November 12, 2012. Revised: October 28, 2014

Incentives are calculated based on annual natural gas savings up to 50% of project costs.

- As the largest social housing provider in the country, projects with Toronto Community Housing (TCH) buildings comprised approximately 16 million CCM of the Part 3 2015 results. The successful implementation of the retrofit projects in TCH buildings is largely attributed to an integrated and solutions-based approach tailored for a customer that has very unique needs and objectives. For example, a gas savings volumetric objective was established at the start of the year and a tiered financial incentive structure was proposed to encourage broader building participation. A working group including both Enbridge and TCH met regularly for project updates and to identify solutions to persistent barriers such as collection and compilation of building information and prioritization of energy savings opportunities. Working together to address identified barriers helps to facilitate the development of business cases for capital investments and funding proposals. For example, Enbridge provided technical advice in support of Regent Park's district energy system project. Both parties have recognized the significant value of this approach. As Enbridge continues to extend services to TCH, the Company plans to adopt the same approach for other large low income customer portfolio managers that require flexibility to encourage deeper participation in DSM offers.

- In recent years, Enbridge has invested in project facilitation and technical support services to social housing providers to help elevate the visibility of energy conservation and encourage energy management practices. One initiative, *Audit to Action* is an audit offer managed by Housing Services Corp. (HSC) and extended to social housing providers and service managers. The energy audits are free to participants with a commitment for implementation of some or all of the measures identified through the audit. Participants to this offer are carefully selected by HSC and service managers to ensure follow-through on audit recommendations. Importantly, the audit report informs business cases for capital investments, calls for funding proposals and/or funding subsidies. There were 14 *Audit to Action* participants in 2014 that went on the implement projects in 2015.

- Marketing efforts to increase awareness of the availability of Low Income offers in privately-owned buildings continued in 2015. The Federation of Rental Housing Providers of Ontario (FRPO) remains the primary industry channel for promoting this work. Joint promotional activities that focused on building owner and resident education and engagement were planned throughout the year. Enbridge is also an active member of FRPO's steering committee for the development of the Environmental Champion's module of FRPO's Certified Rental Buildings (CRB) Program.

- The City of Toronto has been an invaluable partner in cross-promoting the Enbridge Low Income offers to privately-owned multi-residential building managers along with the City's own programs including Tower Renewal and the Local Improvement Charge (LIC) Financing Program. Municipal partnerships such as this provide a template for how the Company will approach other municipalities and focus expansion efforts to private multi-residential buildings in other regions of the Enbridge service territory.

- The Company dedicates significant efforts in strategic outreach to its stakeholders and key customers. These efforts highlight the value of energy management and the Enbridge partnership, while at the same time help to inform the Company's programming activities. Enbridge recognizes the importance of ensuring its offers are providing value-add to the customer's housing operations and are responsive to the changing needs of these customers.

- Enbridge has learned that resident engagement has become a significant factor influencing decision-making within the affordable housing building community. Successful project implementation requires that the Company continue to co-ordinate its efforts with the understanding that resident input to the budgeting considerations and project prioritization of housing providers is part of the process toward project execution.

- The Company plans to continue to drive Part 3 results in 2016 by focusing on the needs of housing providers' and being responsive to the evolving affordable housing landscape.

Low Income Building Performance Management (LIBPM) [19]

Objectives	This offer is designed to provide participants with detailed energy and water consumption information and benchmarking reports at no cost. The goal is to raise the level of awareness on energy usage. In addition, coaching is provided on possible areas of improvement, energy efficiency tips and energy efficiency opportunities.
Target Customer	This offer targets multi-residential social housing providers and managers as well as eligible owners and property managers of privately-owned multi-unit residential buildings (MURBs), which provide housing to a market that includes low income customers and families based on screening criteria established in collaboration with Enbridge's Low Income Consultative Working Group.
Metrics	The metric for this offer is based on the percentage of Part 3 buildings enrolled in the current year. The formula for calculating the percentage of Part 3 buildings enrolled in the current-year Low Income Building Performance Management offer is as follows: $\% \text{ LIBPM} = \frac{(x + y)}{(x + y + z)}$ where: x = Number of new LIBPM buildings in the current year that have participated in another aspect of the Low Income program in a previous year of the 2012-2014 plan; y = Number of new LIBPM buildings participating in current year that have not previously participated in the Low Income program; and, z = Number of buildings in the current year that have implemented custom projects other than LIBPM.
Tracking Methodology	Participating buildings are required to complete an Enrollment and Participation form. Monthly monitoring and tracking is

19 Low Income Building Performance Management is the Low Income offering complement to the Commercial Run it Right (RiR) offering.

	conducted by a third party agent and quarterly reporting is provided to the customer and Enbridge.
Offer Description	As a result of the 2015 rollover, this offer continued as outlined in the 2013-2014 Update (EB-2012-0394). In recognition of the need for a Building Performance Management offer directed at the Low Income sector, the concept of the Commercial Run it Right activity was modified to reflect the needs of social housing providers and the characteristics of social housing buildings. The Low Income Building Performance Management initiative (LIBPM) has been simplified to include: • benchmarking specifically developed for the social housing sector; • analysis of historical consumption data; • development of recommendations for reducing consumption; and • assessment of resulting changes in consumption 12 months later based on changes in actual gas usage. In line with the Low Income delivery strategy of leveraging and/or enhancing existing sector and delivery agents' networks, Enbridge entered into an agreement with HSC to reach social housing buildings. Through this initiative, the energy consumption of participating buildings is tracked over a twelve-month period. Quarterly reports are generated for each building. Follow-up calls are made by HSC to "underperformers" based on the benchmarks established, to provide coaching and identify pathways to energy savings – from improved operational practices to energy savings incentives. The quarterly report is also used to generate program awareness and to identify potential projects for custom or prescriptive offers. In the case of qualified privately owned multi-residential low income buildings, participants were enrolled in Energy Compass

	and benefited from the consumption analysis provided through that initiative.
2015 Results	Enbridge was able to reach a significant number of buildings for participation in benchmarking efforts, with 121 properties that enrolled and participated in 2015. Based on the calculation outlined for the metric, this resulted in a score of 64.7% for this metric, above the upper target for this initiative in the 2015 DSM rollover scorecard.

Offer Commentary and Lessons Learned:

- Housing Services Corporation continued its role as program implementation agent for Enbridge's LIBPM offer through its Utility Management Program. As a sector organization, it plays an important role as a stakeholder and communication channel for the Company's Low Income Program.

- This initiative has been well-received specifically by social housing providers and their service managers. Moving forward, though there is no longer a defined metric included in the scorecard for 2016 and beyond for this type of offer, the practice of benchmarking building performance has become a best practice in good energy management efforts and will continue to be an important facet of Enbridge's engagement with the Multi-Residential Low Income market as an enabling activity to support other offers that will continue in the 2015-2020 Multi-Year DSM Plan.

7. Market Transformation Scorecards

Market Transformation programs are designed with the aim of influencing consumer behaviour and attitudes in support of reducing energy consumption. Market Transformation activities focus on enabling fundamental changes that lead to increased market share of energy efficient products and services, and on influencing consumer behaviour and attitudes that support reductions in natural gas consumption.

Enbridge's Market Transformation program is comprised of two offers which are directed to the new construction sector, both Commercial and Residential, as well as an offer aimed at the existing residential housing sector. As 2015 is a rollover year, these three offers are continuations of offers established in the 2012-2014 Multi-Year DSM Plan.

Enbridge is pleased to report that 2015 was another successful year with respect to the performance of the Market Transformation (MT) program. Efforts in 2015 have focused on continuing to build awareness and recognition in the marketplace, with the aim of educating and influencing the respective target market groups in support of reductions in natural gas consumption.

Introduced in 2012, Savings by Design Residential and Savings by Design Commercial are designed to influence builders and developers in the new construction sector. These offers were developed to provide a basis, both through education and influence, to engage with stakeholders through an interactive assessment process with a focus of exploring design options and construction considerations to construct to standards above building code requirements and achieve energy performance savings.

The Home Labelling (Rating) offer was developed for the home re-sale marketplace and was intended to help educate the realtor community about what a home rating represents and the value it brings to homebuyers and sellers.

Performance in the Market Transformation program is assessed in terms of metrics specific to each of the three offers. On a weighted scorecard basis, all three of the offers exceeded their respective upper performance targets in 2015.

Both the Savings by Design Residential and Commercial offers will continue to be delivered as part of the Market Transformation program in 2016. The Home Labelling offer however, will not continue. Details regarding individual offers are discussed below.

7.1 Residential Savings by Design (SBD)

Objectives	The goal of the Residential Savings by Design offer is to use the Integrated Design Process (IDP) to demonstrate to builders the potential for achieving higher levels of energy and environmental performance through the application of alternative design approaches. In order to realize the potential that the IDP demonstrates to the builder, performance incentives are provided. These incentives encourage the construction of new homes to an energy efficiency standard 25% above the level prescribed in the 2012 Ontario Building Code, ("OBC"). Enbridge expects that Residential SBD will help builders see the value of the IDP approach, and encourage adoption on an ongoing basis.
Target Customer	The offer targets builders and designers of new, Part 9 residential low-rise houses (towns, semis and detached homes) in the Enbridge franchise territory. The intent is to engage builders who construct multiple homes in any given year.
Metrics	There were two metrics for SBD Residential in 2015. The first metric tracks the number of previously non-participating eligible builders that enroll and take part in the IDP; the second metric tracks the number of homes built to the SBD specifications over the course of the year.
Tracking Methodology	This offer requires a commitment from builders to construct homes within a three-year time frame following the completion of the IDP. Commitment letters and eligibility documents along with IDP reports are maintained for all participants. Third party reporting of energy audits is compiled and tracked to support incentive payments. Given the three-year window, in order to follow-up on the builder commitment, Channel Consultants maintain regular contact with

	builders to ensure that all required documentation is provided and proper submission procedures are followed for the builders to receive incentives.
Offer Description	SBD Residential focuses on engaging building industry stakeholders and leveraging industry capabilities to encourage builders to make informed decisions that can realize potential energy savings. Through educating builders on how to construct more energy efficient houses, along with providing a building incentive, the Company influences these builders to first “design it right”, then “build it right” and, finally, “sell it right”. SBD Residential is designed to provide a variety of support activities for builders of new homes from the early design phase through to construction. Savings by Design is a process-based approach involving: • Visioning Session – to define the builder’s sustainability priorities and opportunities; • Integrated Design Process Session – to identify and evaluate strategies to meet the builder’s sustainability goals and the SBD energy reduction target of 25% beyond code through application of energy modelling; • Building Energy Modelling – to evaluate energy performance baselines and proposed improvements. This SBD consultation process involves connecting participating design teams with leading industry experts and other stakeholders as they consider alternative approaches to energy and environmental performance. Through this process, the team works with the builder to explore opportunities to achieve higher energy performance. Starting with the building envelope (windows, wall structure, insulation) and moving inward with HVAC mechanicals and lighting, the

	Savings by Design team guides the builder through a design process to achieve a modelled building that performs to at least 25% better than 2012 OBC. In addition, depending on the specific priorities identified during the visioning session, experts from fields such as lighting, storm water management, sustainable land-use planning, indoor air quality and renewable energy can be engaged to provide further value to the IDP. A third-party service provider undertakes testing and verification to ensure that constructed homes are built with 25% greater energy efficiency than required under the current OBC.
2015 Results	As illustrated in Table 7.0, Residential SBD was successful in enrolling 19 builders who completed the IDP process in 2015. The result exceeds the middle (100%) target for this metric. In addition, there were 1,987 new homes built in relation to the completed units metric. In other words, for builders who have enrolled and completed the IDP process since 2012, there were 1,987 new homes constructed in 2015 through this initiative with features consistent with SBD standards of 25% above OBC (as illustrated in the builder's IDP). This result exceeded the upper target for completed units in 2015.

Table 7.0 2015 Residential Savings by Design Scorecard

Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Residential Savings by Design	*Builders Enrolled* [1]	***60%***	13	18	22	**19**
	Completed Units	***40%***	833	1,111	1,389	**1,987**

1. Eligible builders based on a minimum of 50 homes built in the prior year.

2015 Commentary and Lessons Learned:

- In 2015, SBD Residential saw a slight decrease in participating builders in comparison with 2014, as over the previous 3 years many of the large production builders had already been through the IDP portion of the offer.

- SBD continues to rely on the development of relationships by Channel Consultants with key decision makers in the builder community in order to achieve targets. Channel Consultants have built and continue to build these relationships to encourage builders to reassess their approach to building design as it relates to energy efficiency considerations.

- Enbridge remains strategically involved throughout the builder community, participating actively in builder conferences, education forums and industry associations that provide an opportunity for builder (and other primary stakeholder) engagement and energy efficiency advocacy. For example:
  - Enbridge has representatives involved in the various Home Building Associations across the franchise, provincially and nationally;

- Enbridge representatives sit on the board of directors for BILD, the Sustainable Buildings Canada board, as well as the Canadian Home Builders' Association (CHBA) net zero council.

➢ Over the past 5 years the combination of more stringent mortgage lending rules, steadily increasing housing prices, and increased household debt have made customers more cost conscious when making home purchasing decisions. In addition, builders have expressed growing concerns with increasing development costs and land availability. Notably in 2015, discussions taking place during IDP sessions focused on the need to look for more cost-effective and energy efficient ways to build new homes.

➢ These aforementioned consumer market conditions further support the SBD approach of engaging builders in a "push" strategy to increase energy efficient new home construction. This is not to suggest that the Company discounts the importance of working with builders and other stakeholders to increase awareness and education of energy efficiency in the consumer market, but due to competing consumer interests, consumer demand alone will not drive the changes needed to move the market towards greater levels of energy efficient construction.

➢ Enbridge continues to respond to builder needs addressing the sales and marketing challenges facing the new construction market. The marketing support package that the Company created to support builders in their model homes was well received. Several builders took advantage of the offer to order and use the SBD materials to help promote energy efficiency to potential home buyers. In addition, the IDP optional sales and marketing module has been selected by many participants, and feedback has been positive.

➢ Enbridge participants on the IESO Business and Residential working groups, which includes representatives from IESO as well as LDCs. This served to support Enbridge's efforts in continuing to foster collaboration between CDM and DSM offers as it relates to new construction programming. Consequently, Enbridge was able to provide SBD participants with information on CDM incentives, and this involvement has also provided a forum for planning discussions around future potential New Construction collaboration between DSM and CDM programs.

- In the course of ongoing assessment of the offer and how to best engage and influence builders, stakeholder consultation included:
  - Municipalities – specifically as it related to the support the SBD offer could provide, to help communities meet efficiency objectives and in the execution of Municipal Energy Plans
  - Conservation Authorities
  - Other Industry participants including Energy Modelers, Service Organizations, NRCAN etc.

- In addition, builders that had participated in SBD in previous years have expressed an interest in re-engaging with Enbridge and its team of experts to participate in additional IDPs for new, upcoming developments. As builders typically construct many different designs of homes in different degree day zones, with multiple model variations in response to changing market needs, it would be beneficial to participate in additional IDP's to consider different projects with different challenges. To reflect this need, the offer has been revised in the 2015-2020 Multi-Year DSM Plan.

7.2 Commercial Savings by Design (SBD)

Objectives	The goal of the Commercial Savings by Design offer is to use the Integrated Design Process to demonstrate to builders the potential for achieving higher levels of energy and environmental performance through the application of alternative design approaches. The offer is intended to support this demonstration and awareness with incentives that encourage builders to use the knowledge gained in the IDP to design and build buildings that are more energy efficient. Enbridge expects that Commercial SBD will help builders see the value of the IDP approach, and encourage adoption on an ongoing basis.
Target Customer	This offer is targeted at builders and designers of new, Part 3 commercial buildings in the Enbridge franchise territory. Enbridge targets its promotional activity to owners, builders and developers, design teams including architects, design engineers and energy modelers.
Metrics	Builders and developers who enroll in the offer and complete the IDP process are eligible to be counted towards performance targets. Metrics are based on the number of projects to which a developer commits. As per EB-2012-0394, “the same developer with different clients and different kinds of projects may be counted multiple times. A minimum 100,000 square feet requirement applies to each project. A project is defined as either a single building or multiples of the same building by the same company that adds up to 100,000 square feet.”[20]
Tracking Methodology	Enrollment entails a signed memorandum of understanding with a builder or developer containing a commitment to participate in the Commercial Savings by Design offer and participate in the IDP process. The builder commits to

20 EB-2012-0394, Exhibit B, Tab 1, Schedule 3, page 17 of 20.

	constructing building(s) to the IDP standard within five years in order to receive performance incentives. Enbridge Channel Consultants maintain regular contact with builders to track project status to project completion. Charrette reports for each IDP are maintained to provide a record of information on preliminary estimated savings for each project. All documentation and incentives are tracked by Tracking and Reporting.
Offer Description	Enbridge has provided commercial new construction programming since 1999, beginning with the Design Assistance Program ("DAP"), which was developed to engage the new building design community to design and model new construction buildings to higher levels of energy efficiency. The Commercial Savings by Design offer was designed and developed for delivery beginning in 2012 to encourage developers to build/construct Part 3 buildings to 25% above 2012 OBC. The offer includes the following types of activities: • Improving sizing and design; • Optimization of passive solar, day lighting and natural ventilation; • Integration of high efficiency lighting and HVAC systems; • Integration of lighting and HVAC controls in response to occupant loads; • Reduction and/or optimization of internal loads; • Improving thermal characteristics of the building envelope; and, • Managing environmental impacts. In addition to the facilitation of the IDP, which brings together industry experts, conservation authorities, and municipalities, the offer provides incentives that include financial support to cover costs associated with the IDP and additional incentives tied to the achievement of gas savings above code.

2015 Results	Enbridge was successful in enrolling 24 new developments in 2015 that met the eligibility requirements and completed the IDP process. This result reached the upper scorecard target.

Table 7.1 2015 Commercial Savings by Design Scorecard

Component	Metric	Weight	Targets			2015 Result
			Lower	Middle	Upper	
Commercial Savings by Design	*New Developments Enrolled*	***100%***	11	18	24	**24**

2015 Commentary and Lessons Learned:

- In 2015, Enbridge continued to increase participant levels over previous years as builder interest in the offer remained strong. This can be attributed to an increased awareness of the offer in the market and a better appreciation of the value of participation.

- As the Company continues to learn from participants, the tools employed to market the offer have evolved. In 2015 further enhancements were made based on solicited builder feedback, for example:
  - refinements to content in existing point of sale material were made to better link benefits to barriers faced by builders;
  - a promotional video was created for use at Company sponsored events;
  - additional builder testimonial videos were produced; and,
  - whitepapers and advertorials were published in print media publications.

- While awareness of SBD has increased over the past 3 years, engaging a builder at the right time remains challenging and crucial to securing participation. To that end, Enbridge continues to remain strategically involved throughout the builder community, actively participating in conferences and industry associations that provide a forum for builder (and other primary stakeholder) engagement.

- Specifically Enbridge has been active in the following areas:
  - Canada Green Building Council (Toronto Chapter) as both a sponsor and active participant and presenter at their various events;
  - Sustainable Buildings Canada Board member, actively supporting the planning and execution of the well-attended Green Buildings Festival annual conference; and,
  - supporting municipal energy planning where Enbridge has increased its engagement with municipalities and regions to ensure the SBD offer remains top of mind in examining strategies to meet community efficiency objectives.

- Builders/developers continue to regard energy efficiency as a cost rather than an investment, as their primary objective is to simply meet code requirements on time and on budget. Market price sensitivity for both multi-residential and other commercial building types remains a primary focus for builders, conflicting with an interest to drive down ongoing operational costs that would result from energy efficiency improvements.

- An additional facet to the offer was introduced in 2015; a costing specialist was added to the IDP expert panel. While the estimations of efficiency costs that are provided represent a range of costs that are completely dependent on the incremental energy efficiency approaches selected, builders have responded well to receiving this information as it is helpful in the management of life cycle cost expectations. The additional information is expected to provide a greater likelihood that energy efficiency targets established in the IDP process can be achieved.

- While new condominiums continue to represent a significant percentage of annual commercial new construction starts, Enbridge has had moderate success in enrolling developments which reflect other building types in 2015.

- In 2014, Enbridge identified it would be focusing on long term care, healthcare and school projects following fund approvals by the respective ministries for new construction in these sectors. This targeting strategy proved to be successful. Developments included in the 2015 SBD Commercial offer represented a wide variety of commercial building types, including:
  - Condos,
  - Schools,
  - Offices,
  - Churches,
  - Hospitals
  - Long Term Care
  - 6 story wood construction (mixed use residential/retail) [21]

- In efforts to continue to broaden the impact the SBD Commercial offer can have on the commercial new construction market, Enbridge submitted a revised offer in the 2015-2020 Multi-Year DSM Plan (EB-2015-0049) which outlined a reduced minimum square footage eligibility criteria of 50,000 ft^2 beginning in 2016.

- Also beginning in 2016, in response to builder feedback on performance incentives, the incentive structure has changed to better support builder activities

21 An update to Ontario building code in January 2015 allowed for the construction of 6 story wood buildings

and increased the likelihood that the builder constructs to at least the targeted 25% above code.

Build SAVINGS BY DESIGN™

Savings by Design Commercial
How to Access Incentives

The Savings by Design (SBD) program is for the new construction sector which includes commercial/institutional and multi-unit residential buildings that fall within the scope of Part 3 of the Building Code and have a gross floor area of 100,000 square feet or greater.[1] The SBD Commercial program supports and rewards you through the three primary stages of the project:

1. Integrated Design Process (IDP) Benefit

This one-day IDP workshop is to explore changes to the building design to increase the annual energy performance to a level that is at least **25% better than the Ontario Building Code.**[2] The incentive works as follows:

- The following costs for the IDP, estimated at $25,000, are included: the venue, meals and breaks, Sustainable Buildings Canada (SBC) IDP facilitation team, their experts and support staff, Enbridge staff, real-time energy modelling services and final report.
- All costs associated with the attendance of the members of your team, including designers, modellers and experts, are not covered by Enbridge.

2. Energy Performance Incentive

This $15,000 incentive is paid when the Final Design Stage Energy Model, showing that the energy performance target of at least 25% better than code has been met, **is submitted and approved.** To qualify, you must submit the following:

- The Final Design Stage Building Energy Model and the associated Reference Building Energy Model that follows the energy performance path modelling rules of the selected energy code.
- A Final Design Project Performance Summary, with information taken from the energy models that demonstrates the Enbridge Savings By Design Program Target has been met.
- The Final Design Stage Plans and Specifications for the project, in electronic format.
- An invoice made out to Enbridge in the amount of $15,000 plus HST.

3. Commissioning Incentive

This $15,000 incentive is paid when your Final As-Built and Reference Building Energy Models, and Commissioning Report are submitted and approved. To qualify, you must submit the following:

- The Final As-Built Building Energy Model and the associated Reference Building Energy Model that follows the energy performance path modelling rules of the selected energy code.
- A Project Performance Summary, that demonstrates the Enbridge Savings By Design program target has been met.
- The final As-Built Plans and Specifications.
- The Commissioning Report for the completed project that has been prepared and signed by a commissioning agent who has been accredited by either the Building Commissioning Association (BCA), the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE), or the Association of Energy Engineers (AEE).
- A Confirmation of Implementation of Final Design Energy Efficiency Measures form, signed by a member of the design team having a professional designation provided by either the Ontario Association of Architects, or Professional Engineers Ontario.
- An invoice made out to Enbridge in the amount of $15,000 plus HST.

ENBRIDGE
Life Takes Energy™

- Strategic involvement in conferences and events that provide an opportunity to showcase the offer and market the approach will be continue to be the primary focus for SBD marketing efforts.

- With the contract expiration of the last version of the High Performance New Construction Program (HPNC) CDM program, Enbridge has engaged various LDC’s (Toronto Hydro, Hydro Ottawa, Niagara Peninsula Energy Inc.) as well as the Independent Electricity System Operator (IESO) with the objective to better align future iterations of HPNC with Enbridge’s SBD Commercial New Construction offer.

7.3 Home Labelling (Rating)

Objectives	The objective of the Home Labelling offer is the realization of widespread adoption of a voluntary home labelling system in the residential home resale marketplace. This initiative is aimed at educating the Residential market (both realtors and homeowners) in better understanding the concept of a home energy rating and the value it brings in the resale market. Ultimately, the goal is that a home's energy performance rating becomes a standard condition of sale, similar to home inspections for resale homes.
Target Customer	The immediate target market to support the deployment of a home rating system is realtors and their real estate brokerages. Consequently, collaboration with brokerages willing to commit to promoting Home Labelling and educating real estate agents are key components for effective delivery. The ultimate market is Residential (Rate 1) customers and the real estate agents and brokerages who are listing homes for sale.
Metrics	There are two scorecard performance metrics associated with the Home Labelling offer. The first metric requires Enbridge to secure new commitments from realtors collectively responsible for more than 5,000 (middle target) or 10,000 (upper target) home listings per year. A second metric counts the number of ratings performed by buyers and/or sellers. The rating must either be included in a listing or related marketing materials by the seller or made a condition of sale by the buyer.

Tracking Methodology	Commitment letters from new realtors and home ratings included in Multiple Listing Service (MLS) listings or related marketing materials are tracked and recorded in respect of each of the two metrics.
Offer Description	The Home Labelling offer is designed for participants in the re-sale marketplace and aims at helping consumers understand what a home rating represents and the value it can provide to both homebuyers and purchasers at the time of sale or purchase. The offer also aims to motivate realtors to include energy ratings in marketing material (e.g., MLS).
2015 Results	In 2015, 10 new brokerages committed to participate. As illustrated in Table 7.2, these brokerages are collectively responsible for 41,650 home listings. This result exceeded the upper target established for this metric. The number of recorded home ratings marketed in 2015 was 333. This result fell short of the lower target for the second metric specified for this offer.

Table 7.2 2015 Home Labelling Scorecard

Component	Metric	Weight	Targets			2015 Result
			Lower	Middle	Upper	
Home Labelling	*Realtor Commitments* [1, 2]	***50%***	N/A	5,000	10,000	**41,650**
	Ratings performed	***50%***	2,250	4,500	6,750	**333**

1. Commitments to make provision for a data field to show home energy ratings for all homes listed by participating realtors (industry-wide commitment to include such a field on MLS or similar listing service and/or realtors' commitment to do so with all the homes they list on their own websites, handouts and other consumer material).

2. Commitment from realtors collectively responsible for more than 5,000 (middle target) or 10,000 (upper target) listings/year.

2015 Commentary and Lessons Learned:

- Back in 2009, the Green Energy and Green Economy Act included a proposal to mandate a home labelling system for all re-sale homes in Ontario, however implementation did not follow. With continued anticipated opposition from realtors to a government-enforced program, a voluntary system designed to gain acceptance in the marketplace precipitated the inclusion of a home rating offer in the Company's 2012-2014 DSM plan. The approach was intended to leverage the existing infrastructure to achieve voluntary adoption of getting home ratings completed as a standard practice in much the same way as an offer to purchase a home is made under the provision of a home inspection.

- In the 2015 rollover year, activities continued to focus on securing commitments from brokerages; creating awareness and educating realtors on the value of home energy ratings. Participation in conferences and events supporting the realtor community continued to be a fitting venue for promoting awareness of the Home Labelling initiative. In particular, Enbridge Channel Consultants participated in the annual Realtor Quest conference in Toronto – the largest gathering of Real Estate Board members. These conferences provided an excellent forum for Enbridge to engage with industry stakeholders to promote the offer and schedule follow-up sessions with brokerages to explain the offer parameters and incentives, quantify the value of the offer and the benefits to potential buyers and/or sellers, as well as provide education and training workshops.

- Enbridge has had success with the offer to date as the Company has demonstrated good results in influencing brokerages to commit and gaining realtor attendance at brokerage meetings; however, the Company has not seen the anticipated actual number of home listings with the energy rating promoted.

- Though Enbridge has learned that most home buyers agreed that they value the importance of purchasing an energy efficient home, these same buyers most often do not enquire, nor expect that a house has been energy labelled or rated.

- Enbridge has identified some challenges regarding the adoption of home labelling. The Company has learned there are certainly concerns from both the realtor and legal side that introducing such considerations may delay or complicate expediting the closing of the home sale – parties involved in the transaction generally don't understand what the energy rating is. Also, there is a public perception that energy labels are confusing and don't necessarily depict true operating costs.

- In the 2015-2020 Multi-Year DSM Plan (EB-2015-0049), Enbridge proposed a revised offer beginning in 2016 in which the Company would refocus on the home buyer with efforts to promote energy audits as a means to educate and encourage consumers to have home ratings conducted.

- In the January 20th, 2016 Decision and Order on Enbridge's 2015-2020 Multi-Year DSM Plan (EB-2015-0049), the OEB did not approve the continuation of the Home Rating program as part of Enbridge's DSM portfolio beginning in 2016.

- Though the Company will not be proceeding with the Home Labelling offer, Enbridge continues to support the value to consumers of getting an energy audit completed on their homes. The completion of a pre and post-retrofit energy evaluation continues to be a key component of the Home Energy Conservation offer and provides the participant with an Energy Rating score. For those Ontarians increasingly interested in looking for ways to conserve energy and make environmentally responsible choices, working towards and demonstrating a good energy score will not only provide comfort and peace of mind, but will also add value to their homes.

- Enbridge worked with the Ministry of Energy during 2015 to provide input on a proposal being developed to require home energy rating and disclosure (HER&D) at the time of listing. The Company has provided feedback based on experience in the past number of years; explaining the gaps, barriers and successes in an effort to assist the Province in the delivery of a program that consumers can understand.

- The Company will continue to monitor developments at the government level regarding the implementation of a mandated home rating framework and engage with stakeholders where appropriate to provide feedback and support implementation.

8. Lost Revenue Adjustment Mechanism (LRAM)

The LRAM is a mechanism to adjust for margins the utility loses/gains if its DSM program is more/less successful in the period after rates are set than was planned in setting the rates. As outlined in the Guidelines, the LRAM Variance Account (LRAMVA) is used to track, by rate class, the impact of DSM activities undertaken in relation to the forecasted impact included in distribution rates.

LRAM is calculated on a monthly basis using the volumetric impact of the measures implemented. The LRAM amount is an adjustment which may be an amount refundable to, or receivable from, the Company's customers (depending on whether the actual natural gas savings resulting from the natural gas utility's DSM activities are less than or greater than what was included in the forecast for rate-setting purposes).

Table 8.0 2015 LRAM Statement

2015 Annual Report LRAM Calculation

Based on 57,036,910 FE m3 built into rates

Rate Class	Budget Net Partially Effective	Actual Net Partially Effective	Volume Variance	Distribution Margin	LRAM Allocation $	Actual LRAM $
Rate 110	2,065,678	1,254,638	(811,041)	1.4924	($12,104)	$18,724
Rate 115	1,314,523	813,986	(500,536)	0.8174	($4,092)	$6,654
Rate 135	0	25,393	25,393	1.2825	$326	$326
Rate 145	2,428,288	153,892	(2,274,396)	1.5224	($34,626)	$2,343
Rate 170	4,942,907	206,432	(4,736,475)	0.4504	($21,333)	$930
Totals	10,751,396	2,454,340	-8,297,056		($71,829)	$28,976
		Amount to be paid back to Ratepayers			($71,829)	

** Rate 1 and Rate 6 are not included in the LRAM amount for clearance above as these rate classes are covered under the Average Use True-Up Variance Account (AUTUVA)*

9. DSM Shareholder Incentive

Enbridge earns a shareholder incentive based on its performance against targets outlined for Resource Acquisition, Low Income and Market Transformation scorecards. The DSM Incentive provides that incentive to the Company in relation to its DSM activities. Further to approved amounts outlined in EB-2015-0049, Table 9.0 summarizes how the maximum incentive available in 2015 is allocated across each program.

Table 9.0 2015 DSM Maximum Incentive Allocation

Program	Total Budget *	% of Total	Maximum Incentive Available
Resource Acquisition	$19,175,275	58%	$6,482,744
Low Income	$7,382,078	23%	$2,495,721
Market Transformation	$6,244,587	19%	$2,111,159
	$32,801,939 *	**100%**	**$11,089,624**

* The Total Budget reflects the OEB approved 2015 roll-over budget amount of $32,801,939 and does not include the additional Board approved incremental budget of $4,920,291. The combined approved Total DSM budget for 2015 is $37,722,230

Scorecard results and the corresponding DSMI earned for each program is detailed in the following tables:

Table 9.1 Resource Acquisition Scorecard Achievement & DSMI

Resource Acquisition						
Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Volumes	*Cumulative Savings (million m^3)*	**92%**	758.9	1011.9	1264.9	**734.13**
Residential Deep Savings	*Number of Participants*	**8%**	571	762	952	**5,646**
		Total Weighted Scorecard Target Achieved				**150.0%** *
		Scorecard Incentive Achieved				**$6,482,744**

* Weighted scorecard is capped at 150%. Actual scorecard achievement is 152.3%

Table 9.2 Low Income Scorecard Achievement & DSMI

Low Income						
Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Single Family (Part 9)	*Cumulative Savings (million m^3)*	50%	18.1	24.1	30.2	28.07
Multi-residential (Part 3)	*Cumulative Savings (million m^3)*	45%	51.6	68.7	86.0	63.97
Multi-residential (Part 3) LIBPM	*Percent of Part 3 Participants Enrolled*	5%	30%	40%	50%	65%
		Total Weighted Scorecard Target Achieved				116.2%
		Scorecard Incentive Achieved				$1,483,792

Table 9.3 Market Transformation – Residential SBD Scorecard Achievement & DSMI

Market Transformation						
Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Residential Savings by Design	*Builders Enrolled*	60%	13	18	22	19
	Completed Units	40%	833	1,111	1,389	1987
		Total Weighted Scorecard Target Achieved				150.0% *
		Scorecard Incentive Achieved				$1,076,493

* Weighted scorecard is capped at 150%. Actual scorecard achievement is 170.5%

Table 9.4 Market Transformation – Commercial SBD Scorecard Achievement & DSMI

Market Transformation						
Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Commercial Savings by Design	*New Developments Enrolled*	100%	11	18	24	24
		Total Weighted Scorecard Target Achieved				150.0%
		Scorecard Incentive Achieved				$418,269

Table 9.5 Market Transformation – Home Labelling Scorecard & DSMI

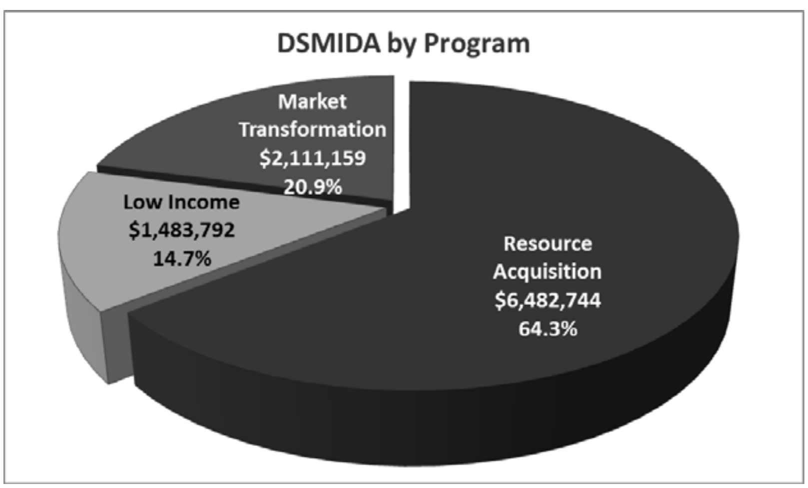
Market Transformation						
Component	Metric	Weight	Targets Lower	Targets Middle	Targets Upper	2015 Result
Home Labelling	*Realtor Committments*	**50%**	N/A	5,001	10,001	**41,650**
	Ratings performed	**50%**	2250	4,500	6,750	**333**
		Total Weighted Scorecard Target Achieved				**150.0%** *
		Scorecard Incentive Achieved				**$616,397**

* Weighted scorecard is capped at 150%. Actual scorecard achievement is 237%

Table 9.6 2015 DSMIDA Summary

Program	DSMIDA by Program	% Contribution
Resource Acquisition	$6,482,744	64%
Low Income	$1,483,792	15%
Market Transformation	$2,111,159	21%
TOTAL	**$10,077,695**	**100%**

Table 9.7 2015 Program Contribution to DSMIDA



10. 2015 Budget and Program Spending

10.1 Budget

Table 10.0 provides the 2015 DSM budget as outlined in the 2015-2020 Multi-Year DSM Plan (EB-2015-0049). The Board approved a 2015 budget of $37,722,230 in its Decision on January 20th, 2016.

Table 10.0 2015 DSM Plan Budget

Program	Program Budget	Overheads	Total Budget
Resources Acquisition	$14,443,790	$4,731,485	$19,175,275
Low Income	$6,864,090	$517,988	$7,382,078
Market Transformation	$4,890,900	$1,353,687	$6,244,587
Total Roll-over Budget	**$26,198,780**	**$6,603,160**	**$32,801,939**
Incremental Budget			$4,920,291
Total 2015 DSM Budget			**$37,722,230**

10.2 2015 Spending

Table 10.1 2015 OEB Approved Budget vs. Spending

Program	OEB Approved Budget	2015 Spending	Variance	%
Resource Acquisition	**$19,175,275**	**$23,389,805**	**$4,214,530**	**22%**
Residential	*$1,872,720*	*$9,362,295*	*$7,489,575*	
Commercial	*$8,252,370*	*$6,221,724*	*-$2,030,646*	
Industrial	*$4,318,700*	*$2,166,706*	*-$2,151,994*	
Overheads	*$4,731,485*	*$5,639,080*	*$907,595*	
Low Income	**$7,382,078**	**$7,173,710**	**-$208,368**	**-3%**
Part 9 Residential	*$4,655,790*	*$4,444,616*	*-$211,174*	
Part 3 Multi-Residential	*$2,208,300*	*$2,111,746*	*-$96,554*	
Overheads	*$517,988*	*$617,349*	*$99,361*	
Market Transformation	**$6,244,587**	**$4,657,079**	**-$1,587,508**	**-25%**
Residential SBD	*$2,493,900*	*$2,032,022*	*-$461,878*	
Commercial SBD	*$969,000*	*$890,464*	*-$78,536*	
Home Labelling	*$1,428,000*	*$121,241*	*-$1,306,759*	
Overheads	*$1,353,687*	*$1,613,352*	*$259,665*	
Total	**$32,801,939**	**$35,220,594**	**$2,418,654**	**7%**
Incremental Budget	**$4,920,291**	**$559,378**	**-$4,360,913**	
Grand Total	**$37,722,230**	**$35,779,972**	**-$1,942,258**	**-5%**

As outlined in Table 10.1 above, spending in relation to Enbridge's DSM programming in 2015 was $35,220,594. In addition $559,378 was spent against incremental budget initiatives (discussed in further detail below). Total spending amounted to $35,779,972.

10.3 Incremental Budget

In its original application for the 2015-2020 Multi-Year DSM Plan (EB-2015-0049) filed on April 1st, 2015, Enbridge identified a series of initiatives it believed were appropriate pursuits above and beyond the 2015 rollover budgets which would help to begin transitioning into a new DSM Framework. As outlined in Section 15.1 of the DSM Framework these amounts, not having been built into to rates, will be recovered via the DSMVA and, as stated at page 38 of the Filing Guidelines will be "…incremental to any DSMVA amounts used…" for the purpose of achieving results beyond the 100% achievement level. As noted in Section 2.1 of this report the Board approved an incremental budget of $4.92 million for this purpose in pursuit of the guiding principles and key priorities outlined in the 2015-2020 DSM Framework.

Throughout the course of the EB-2015-0049 proceeding in 2015 approval of Enbridge's overall incremental budget and the items listed within it were not certain. This uncertainty was compounded by a recommendation by a party to the proceeding which Enbridge adopted in its Reply Argument that would allow for any approved incremental budget to be carried forward into 2016 given the merit of the initiatives proposed and the likelihood that uncertainty would persist into the final months of 2015, or even into 2016.

As a result, the Company proceeded cautiously in spending this budget within 2015. Ultimately the Board's Decision and Order in EB-2015-0049 was not received until January 20th, 2016, in which the Board disallowed the carrying forward of the 2015 incremental budget into 2016. The net result of all these factors was significant underspending on all items listed within the incremental budget.

Table 10.3 below outlines each item within the incremental budget inclusive of a basic description of the initiative, the budget approved by the Board in EB-2015-0049, the actual spending on each initiative within 2015, and a brief explanation of the variance for each item.

Table 10.2 Incremental Budget vs. 2015 Incremental Spending

Incremental Budget Item	Incremental Budget	2015 Spending
My Home Health Record (MMHR)	*$2,650,000*	*$444,801*
Integrated Resource Planning	*$300,000*	*$0*
Potential Study Update	*$50,000*	*$0*
Green Button Initiative	*$300,000*	*$0*
Comprehensive Energy Management	*$370,000*	*$60,462*
Low Income New Construction	*$250,000*	*$1,101*
Collaboration Fund	*$1,000,000*	*$53,014*
Unallocated	*$291*	*$0*
TOTAL	**$4,920,291**	**$559,378**

Table 10.3 Incremental Spending Detail

Budget Item	Description	Approved Budget (000's)	Actual Spending (000's)	Explanation of Variance
My Home Health Record Residential Behaviour Program(MHHR)	Rollout of the MHHR offer to residential customers in the first partial year of the offer.	$2,650.0	$444.8	While the Board's Decision approved Enbridge's incremental budget of $4.92M which included MHHR costs in 2015, it did not approve continuation of MHHR into 2016. Spending was incurred prior to the Board's Decision on Jan. 20th, 2016.
Integrated Resource Planning (IRP) Study	Undertaking of the Board's guidance to conduct an IRP study. This study is to be completed in time to inform the mid-term review.	$300.0	$0.0	Enbridge's EB-2015-0049 Application included a proposed scope of work for an IRP study. Given that neither approval of the scope of work nor the incremental budget were received until 2016, Enbridge did not commence spending on an IRP study in 2015
Potential Study Update	Work towards completing an update to recent Potential Study in order to account for and incorporate more recent market potential data that becomes available. And/or contribute towards funding ground up research in collaboration with Union and the IESO to better inform a sector by sector understanding.	$50.0	$0.0	Enbridge initially budgeted this amount to augment its potential study, filed as Exhibit C, Tab 1, Schedule 1 in EB-2015-0049 with additional primary research. After requesting this amount, Enbridge learned that the Board itself would be commissioning a new and separate potential study for completion by June of 2016.

Green Button Initiative	Participate in the Green Button initiative sponsored by the MOE. Including development of a customer information system(s) to allow for data transfer.	$300.0	$0.0	Spending on the Green Button Initiative was not initiated in 2015 because timelines for the initiative were slower than expected and the Company did not receive approval of this amount until after the year was already completed.
Comprehensive Energy Management	Offer Comprehensive Energy Management to large industrial and commercial customers.	$370.0	$60.5	Not having approval of the incremental budget until 2016 Enbridge proceeded cautiously in soft-launching CEM. Despite hesitant spending the Company was able to identify a number of priority customers with high suitability for the offer moving forward.
Low Income New Construction	Initiate Low Income New Construction offer.	$250.0	$1.1	Not having approval of the incremental budget until 2016, activities toward this end in 2015 were limited to the exploration of partnership opportunities for the program in 2016 and beyond.
Collaboration and Innovation Fund (CIF)	Fund for collaborative pilot programs to drive understanding on innovative technologies and market approaches.	$1,000.0	$53.0	Uncertainty with respect to Board Approval guided the Company towards a measured and cautious approach towards spending this budget. The $53k spent was largely focused on the development of future collaborative pilots, research and initiatives as the Company felt it prudent to continue to explore collaboration given the strong indications from the Board to do so.
TOTAL		**$4,920.3**	**$559.4**	

10.4 Collaboration and Innovation Fund

As noted on page 82 of the Board's Decision and Order in EB-2015-0049, the Board has approved Enbridge's proposal for a Collaboration and Innovation Fund ("CIF") of approximately $6 million over the term of the 2015-2020 Multi-Year DSM Plan. The Board made the determination for these funds to be available throughout the term of the Plan, rather than approve a distinct $1 million within each year, in contemplation of the need "…to provide flexibility and address important opportunities when presented."[22]

22 EB-2015-0049, Decision and Order, p.81, January 20th, 2016

As a result of the lack of certainty during 2015 regarding the future of the CIF specifically, and the incremental budget more broadly, spending of this Fund in 2015 was limited. During 2015 approximately $53,000 was spent.

CIF spending in 2015 was largely focused on the development of future collaborative pilots, research and initiatives. While these early efforts did not generate distinct gas saving or other results in 2015, it is anticipated that they will facilitate meaningful collaboration with the Independent Electricity System Operator ("IESO") and electric utilities in 2016 and beyond. Of note, a small commercial and industrial collaborative energy assessment effort was undertaken with Enersource Corporation ("Enersource"). The initiative involved 30 commercial and 20 industrial customer site visits. The purpose of the initiative was to generate energy efficiency awareness and engagement amongst these customer segments, identify opportunities for customers to save both electricity and gas through DSM and CDM programs, gain further insight into the needs of these customer segments, increase customer convenience and also reduce the cost of these activities.

10.5 Demand Side Management Variance Account (DSMVA)

As specified in the Guidelines, the DSMVA "should be used to track the variance between actual DSM spending by rate class versus the budgeted amount included in rates by rate class."[23]

The exact DSM budget built into rates for the 2015 calendar year was $34,954,513. This amount was proposed by Enbridge and subsequently approved by the Board in the Company's 2015 Rate Adjustment proceeding EB-2014-0276. The following excerpt filed November 28th, 2014 outlines the rationale for the inclusion of this amount in rates:

> The Framework Consultation will result in a new DSM Framework that will apply to the six years 2015 through 2020. While the Company and DSM stakeholders recently filed submissions on the Board's draft DSM Framework and Guideline released on September 15, 2015, a final decision of the Board on the new DSM Framework is not expected until later this year. The timing of the Board's decision on the Framework Consultation creates some uncertainty and complexity from a planning perspective for 2015. Despite this, the

23 EB-2014-0134. Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2020), OEB, December 22, 2014, Page 38.

> Company has developed a 2015 DSM budget it believes recognizes directionally the likely result of the Framework Consultation and reflects the practicalities of the planning process for DSM programs that will be operated in 2015. Enbridge is therefore proposing an increase in the DSM Budget for 2015 to $35 million…
> The increase in the filed budgets from $32.8 to $35 million for 2015 results from direction ascertained from two notable milestones on the evolution of DSM beyond the current DSM Framework. These milestones included: 1) the Minister of Energy's Directive to the Board outlining the requirement for a six-year plan and achievement of all cost-effective DSM; and, 2) the Draft DSM Framework and Guidelines released by the Board on September 15, 2014 outlining preliminary guidance on the level of budget for consideration between 2015 and 2020.[24]

Subsequent to submitting the above-noted evidence, Enbridge received the Board's final DSM Framework and accompanying Filing Guidelines on December 22nd, 2014. The final Framework was inclusive of direction to roll-forward Enbridge's 2015 DSM budget and an invitation to apply for an additional 15% in budget to address the Board's Guiding Principles and Key Priorities. As addressed previously in Section 2.1, the combined effect of these two Framework elements resulted in a proposed budget of $37.7 million, which was approved by the Board in its Decision and Order in EB-2015-0049 on January 20th, 2016.

Table 10.4 shows the variance between the 2015 DSM spending (as summarized previously in Table 10.1) in comparison to the DSM budget built into rates. The resulting DSMVA for 2015 is $825,460 (recoverable from ratepayers).

Table 10.4 DSMVA Determination: 2015 Spending vs. Amount Built Into Rates

DSM Budget Previously Built Into Rates	2015 DSM Spending	Spending Against Incremental Budget	Variance = DSMVA
A	B	C	Variance = A-B-C
$34,954,513	**$35,220,594**	**$559,378**	**-$825,460**

24 EB-2014-0276, Exhibit D1, Tab 4, Schedule 1, pages 1-2

10.6 DSM Rate Allocation

Table 10.5 illustrates the allocation to rate classes of the DSM Variance Accounts.[25]

Table 10.5 2015 Rate Allocation

Rate Class	DSMIDA	LRAMVA	DSMVA	TOTAL
Rate 1 [1]	$5,901,877	N/A	$6,498,202	$12,400,079
Rate 6 [1]	$3,438,449	N/A	-$2,919,026	$519,423
Rate 9 [2]	$404	$0	-$74	$331
Rate 100	$0	$0	$0	$0
Rate 110	$421,703	-$12,104	$191,445	$601,044
Rate 115	$146,222	-$4,092	-$451,019	-$308,888
Rate 125 [2]	$15,156	$0	-$2,779	$12,377
Rate 135	$12,883	$326	-$144,351	-$131,142
Rate 145	$41,385	-$34,626	-$1,040,849	-$1,034,090
Rate 170	$93,350	-$21,333	-$1,304,940	-$1,232,923
Rate 200 [2]	$5,254	$0	-$963	$4,291
Rate 300 [2]	$1,010	$0	-$186	$825
Total	**$10,077,695**	**-$71,829**	**$825,460**	**$10,831,326**

1. Rate 1 and Rate 6 are not included in the LRAM amount as these rate classes are covered under the Average Use True-Up Variance Account (AUTUVA).

2. Rates 9, 125, 200 & 300 do not have any LRAM component in the rate allocation since customers in these rate classes are not eligible for DSM programs. These rate classes will however be subject to rate allocations for DSMVA and applicable DSMIDA related to Low Income Program.

25 As in prior years, Low Income DSM spending is allocated to all rate classes, to be consistent with the electricity conservation framework, as well as the LEAP Emergency Financial Assistance program. Allocation for the LEAP fund was outlined in EB-2008-0150 Report of the Board: Low Income Energy Assistance Program on page 11 Section 5.1.1 Funding LEAP.

Appendix A: Input Assumptions

On March 27th, 2015, Enbridge Gas Distribution Inc. and Union Gas Limited submitted a joint application which sought approval from the Ontario Energy Board for new and updated Demand Side Management measures. The Board assigned file number EB-2014-0354 to this application. On July 23rd, 2015 Enbridge and Union Gas were granted approval of the new and updated DSM measures and input assumptions as set out in the joint application, EB-2014-0354.

These inputs were subject to Enbridge’s 2014 DSM audit. The inputs were used in calculating the savings claims that comprise the 2015 DSM scorecard results. The Board approved this approach as per its revised Decision and Order issued February 24th, 2016 (EB-2015-0049).

Appendix B: 2015 Avoided Costs

The following tables outline the Avoided Costs used in the determination of 2015 results and are included here for reference in the following charts:

	2015 Gas Avoided Costs							
	Water Heating		Space Heating		Combined Space & Water Heating		Industrial	
Year	Baseload ($/m3)		Baseload ($/m3)		Baseload ($/m3)		Baseload ($/m3)	
	Rate	NPV	Rate	NPV	Rate	NPV	Rate	NPV
1	0.1309	0.1342	0.1375	0.1495	0.1365	0.1476	0.1312	0.1347
2	0.1689	0.2939	0.1784	0.3182	0.1767	0.3147	0.1700	0.2981
3	0.1846	0.4590	0.1950	0.4926	0.1931	0.4873	0.1857	0.4698
4	0.1974	0.6258	0.2087	0.6690	0.2066	0.6621	0.1987	0.6465
5	0.2120	0.7954	0.2241	0.8482	0.2219	0.8395	0.2134	0.8289
6	0.2249	0.9654	0.2377	1.0279	0.2353	1.0174	0.2263	1.0149
7	0.2356	1.1338	0.2490	1.2060	0.2465	1.1937	0.2371	1.2023
8	0.2403	1.2963	0.2539	1.3777	0.2514	1.3636	0.2418	1.3861
9	0.2320	1.4446	0.2452	1.5344	0.2427	1.5189	0.2335	1.5567
10	0.2257	1.5810	0.2385	1.6787	0.2362	1.6616	0.2272	1.7163
11	0.2302	1.7126	0.2433	1.8178	0.2409	1.7994	0.2317	1.8729
12	0.2348	1.8396	0.2482	1.9519	0.2457	1.9322	0.2363	2.0264
13	0.2395	1.9620	0.2531	2.0814	0.2506	2.0603	0.2411	2.1770
14	0.2443	2.0801	0.2582	2.2062	0.2556	2.1839	0.2459	2.3246
15	0.2492	2.1940	0.2634	2.3266	0.2607	2.3031	0.2508	2.4695
16	0.2541	2.3039	0.2686	2.4427	0.2660	2.4181	0.2558	2.6115
17	0.2592	2.4099	0.2740	2.5547	0.2713	2.5290	0.2609	2.7508
18	0.2644	2.5121	0.2795	2.6628	0.2767	2.6359	0.2662	2.8875
19	0.2697	2.6107	0.2851	2.7670	0.2822	2.7391	0.2715	3.0215
20	0.2751	2.7058	0.2908	2.8675	0.2879	2.8386	0.2769	3.1529
21	0.2806	2.7975	0.2966	2.9644	0.2936	2.9346	0.2824	3.2818
22	0.2862	2.8860	0.3025	3.0579	0.2995	3.0272	0.2881	3.4082
23	0.2919	2.9713	0.3086	3.1481	0.3055	3.1165	0.2939	3.5322
24	0.2978	3.0536	0.3147	3.2351	0.3116	3.2026	0.2997	3.6538
25	0.3037	3.1330	0.3210	3.3190	0.3178	3.2857	0.3057	3.7731
26	0.3098	3.2096	0.3275	3.4000	0.3242	3.3658	0.3118	3.8901
27	0.3160	3.2834	0.3340	3.4780	0.3307	3.4431	0.3181	4.0048
28	0.3223	3.3547	0.3407	3.5533	0.3373	3.5177	0.3244	4.1173
29	0.3288	3.4234	0.3475	3.6260	0.3440	3.5896	0.3309	4.2277
30	0.3353	3.4896	0.3544	3.6960	0.3509	3.6589	0.3376	4.3359

The Nominal Growth Rate used in the table is 2.0%

The Real Discount Rate used in the table is 4%, the Nominal Discount Rate used in the table is 5.75%

2015 Water and Electricity Avoided Costs

Year	Water Heating				Space Heating				Combined Space & Water Heating				Industrial			
	Electricity (¢/Kwh)		Water ($/1000 litre)		Electricity (¢/Kwh)		Water ($/1000 litre)		Electricity (¢/Kwh)		Water ($/1000 litre)		Electricity (¢/Kwh)		Water ($/1000 litre)	
	Rate	NPV	Rate	NPV	Rate	NPV	Rate	NPV	Rate	NPV	Rate	NPV	Rate	NPV	Rate	NPV
1	0.1132	$0.11	0.68875	$0.69	0.1132	$0.11	0.6888	$0.69	0.1132	$0.11	0.6888	$0.69	0.1132	$0.11	0.6888	$0.69
2	0.1157	$0.22	0.70421	$1.35	0.1157	$0.22	0.7042	$1.35	0.1157	$0.22	0.7042	$1.35	0.1157	$0.22	0.7042	$1.35
3	0.1181	$0.33	0.71871	$2.00	0.1181	$0.33	0.7187	$2.00	0.1181	$0.33	0.7187	$2.00	0.1181	$0.33	0.7187	$2.00
4	0.1206	$0.43	0.73352	$2.62	0.1206	$0.43	0.7335	$2.62	0.1206	$0.43	0.7335	$2.62	0.1206	$0.43	0.7335	$2.62
5	0.1230	$0.53	0.74862	$3.22	0.1230	$0.53	0.7486	$3.22	0.1230	$0.53	0.7486	$3.22	0.1230	$0.53	0.7486	$3.22
6	0.1256	$0.62	0.76404	$3.79	0.1256	$0.62	0.7640	$3.79	0.1256	$0.62	0.7640	$3.79	0.1256	$0.62	0.7640	$3.79
7	0.1282	$0.72	0.77978	$4.35	0.1282	$0.72	0.7798	$4.35	0.1282	$0.72	0.7798	$4.35	0.1282	$0.72	0.7798	$4.35
8	0.1308	$0.80	0.79584	$4.89	0.1308	$0.80	0.7958	$4.89	0.1308	$0.80	0.7958	$4.89	0.1308	$0.80	0.7958	$4.89
9	0.1335	$0.89	0.81223	$5.41	0.1335	$0.89	0.8122	$5.41	0.1335	$0.89	0.8122	$5.41	0.1335	$0.89	0.8122	$5.41
10	0.1362	$0.97	0.82896	$5.91	0.1362	$0.97	0.8290	$5.91	0.1362	$0.97	0.8290	$5.91	0.1362	$0.97	0.8290	$5.91
11	0.1391	$1.05	0.84603	$6.39	0.1391	$1.05	0.8460	$6.39	0.1391	$1.05	0.8460	$6.39	0.1391	$1.05	0.8460	$6.39
12	0.1419	$1.13	0.86346	$6.86	0.1419	$1.13	0.8635	$6.86	0.1419	$1.13	0.8635	$6.86	0.1419	$1.13	0.8635	$6.86
13	0.1448	$1.20	0.88124	$7.31	0.1448	$1.20	0.8812	$7.31	0.1448	$1.20	0.8812	$7.31	0.1448	$1.20	0.8812	$7.31
14	0.1478	$1.27	0.89939	$7.75	0.1478	$1.27	0.8994	$7.75	0.1478	$1.27	0.8994	$7.75	0.1478	$1.27	0.8994	$7.75
15	0.1509	$1.34	0.91791	$8.17	0.1509	$1.34	0.9179	$8.17	0.1509	$1.34	0.9179	$8.17	0.1509	$1.34	0.9179	$8.17
16	0.1540	$1.41	0.93682	$8.57	0.1540	$1.41	0.9368	$8.57	0.1540	$1.41	0.9368	$8.57	0.1540	$1.41	0.9368	$8.57
17	0.1571	$1.47	0.95611	$8.96	0.1571	$1.47	0.9561	$8.96	0.1571	$1.47	0.9561	$8.96	0.1571	$1.47	0.9561	$8.96
18	0.1604	$1.53	0.97581	$9.34	0.1604	$1.53	0.9758	$9.34	0.1604	$1.53	0.9758	$9.34	0.1604	$1.53	0.9758	$9.34
19	0.1637	$1.59	0.99590	$9.70	0.1637	$1.59	0.9959	$9.70	0.1637	$1.59	0.9959	$9.70	0.1637	$1.59	0.9959	$9.70
20	0.1671	$1.65	1.01642	$10.05	0.1671	$1.65	1.0164	$10.05	0.1671	$1.65	1.0164	$10.05	0.1671	$1.65	1.0164	$10.05
21	0.1705	$1.71	1.03735	$10.39	0.1705	$1.71	1.0373	$10.39	0.1705	$1.71	1.0373	$10.39	0.1705	$1.71	1.0373	$10.39
22	0.1740	$1.76	1.05872	$10.72	0.1740	$1.76	1.0587	$10.72	0.1740	$1.76	1.0587	$10.72	0.1740	$1.76	1.0587	$10.72
23	0.1776	$1.81	1.08053	$11.04	0.1776	$1.81	1.0805	$11.04	0.1776	$1.81	1.0805	$11.04	0.1776	$1.81	1.0805	$11.04
24	0.1812	$1.86	1.10279	$11.34	0.1812	$1.86	1.1028	$11.34	0.1812	$1.86	1.1028	$11.34	0.1812	$1.86	1.1028	$11.34
25	0.1850	$1.91	1.12550	$11.64	0.1850	$1.91	1.1255	$11.64	0.1850	$1.91	1.1255	$11.64	0.1850	$1.91	1.1255	$11.64
26	0.1888	$1.96	1.14869	$11.92	0.1888	$1.96	1.1487	$11.92	0.1888	$1.96	1.1487	$11.92	0.1888	$1.96	1.1487	$11.92
27	0.1927	$2.00	1.17235	$12.19	0.1927	$2.00	1.1724	$12.19	0.1927	$2.00	1.1724	$12.19	0.1927	$2.00	1.1724	$12.19
28	0.1967	$2.05	1.19650	$12.46	0.1967	$2.05	1.1965	$12.46	0.1967	$2.05	1.1965	$12.46	0.1967	$2.05	1.1965	$12.46
29	0.2007	$2.09	1.22115	$12.71	0.2007	$2.09	1.2212	$12.71	0.2007	$2.09	1.2212	$12.71	0.2007	$2.09	1.2212	$12.71
30	0.2048	$2.13	1.24631	$12.96	0.2048	$2.13	1.2463	$12.96	0.2048	$2.13	1.2463	$12.96	0.2048	$2.13	1.2463	$12.96

The Nominal Growth Rate used in the table is 2.0%

The Real Discount Rate used in the table is 4%, the Nominal Discount Rate used in the table is 5.75%

2015 DSM DEFERRAL AND VARIANCE ACCOUNT

RATE ALLOCATION AND CLEARANCE OF 2015 DSM BALANCES

1. The following evidence describes the three DSM Deferral and Variance Accounts and provides the allocation of the balances to rate classes. As explained in Exhibit A, Tab 1, Schedule 3, these balances reflect all 2015 verified program results with the exception of the incomplete Net-to-Gross ("NTG") study findings.

2. Consistent with the treatment of 2015 as a transition year, Enbridge has rolled-forward its 2014 DSM plan into 2015.[1] For that reason, the allocation of 2015 DSM deferral and variance account balances to rate classes is consistent with the allocation methodologies approved by the OEB in Enbridge's 2014 DSM Clearance of Variance Accounts Application (EB-2015-0267) and as outlined in the previous Guidelines.

3. Demand Side Management Variance Account ("DSMVA"): The DSMVA is the account that should be "used to track the variance between actual DSM spending by rate class versus the budgeted amount included in rates by rate class."[2] The actual DSMVA spending variance amount versus budget targeted to each customer class is allocated to that customer class for rate recovery purposes.

4. Demand Side Management Incentive Deferral Account ("DSMIDA"): The purpose of the DSMIDA is to record the shareholder incentive amount earned by a natural gas utility as a result of its DSM programs.[3] DSM shareholder incentive amounts are

---

[1] EB-2014-0134 Demand Side Management Framework for Natural Gas Distributors (2015-2020), p. 37.
[2] EB-2014-0134, Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2020), p. 38.
[3] Ibid., p. 39.

allocated to the rate classes in proportion to the amount actually spent on each respective rate class.

5. Lost Revenue Adjustment Mechanism Variance Account (“LRAMVA”): The LRAMVA is the account that “should be used to track, at the rate class level, the actual impact of DSM activities undertaken by the natural gas utility from the forecasted impact included in distribution rates.[4] The LRAM amount is recovered in rates on the same basis as the lost revenues were experienced so that the LRAM ends up being a full true-up by rate class.

The following table illustrates the allocation to rate classes of the 2015 DSM Variance Account balances.

[4] Ibid., p. 39.

2015 Rate Allocation

Rate Class	DSMIDA	LRAM	DSMVA	TOTAL
Rate 1*	$5,901,877	N/A	$6,498,202	$12,400,079
Rate 6*	$3,438,449	N/A	-$2,919,026	$519,423
Rate 9**	$404	$0	-$74	$331
Rate 110	$421,703	-$12,104	$191,445	$601,044
Rate 115	$146,222	-$4,092	-$451,019	-$308,888
Rate 125**	$15,156	$0	-$2,779	$12,377
Rate 135	$12,883	$326	-$144,351	-$131,142
Rate 145	$41,385	-$34,626	-$1,040,849	-$1,034,090
Rate 170	$93,350	-$21,333	-$1,304,940	-$1,232,923
Rate 200**	$5,254	$0	-$963	$4,291
Rate 300**	$1,010	$0	-$186	$825
Total	$10,077,695	-$71,829	$825,460	$10,831,326

** Rate 1 and Rate 6 are not included in the LRAM amount for clearance above as these rate classes are covered under the Average Use True-Up Variance Account (AUTUVA).*

*** Rates 9, 125, 200 & 300 will not have any LRAM component included in the rate allocation since customers in these rates classes are not eligible for DSM programs. These rate classes will however, be subject to rate allocations for DSMVA and applicable DSMIDA related to the Low Income Program.*

Note: Numbers may not add up due to rounding.

ESTIMATED TYPICAL BILL IMPACTS BASED ON 2015 DSM VARIANCE ACCOUNT BALANCES RATE ALLOCATION

The table below provides the estimated impact of the Clearance of the 2015 DSM Variance Accounts on a typical customer's bill in each of the rate classes affected.

Rate Class	Annual Volume for Typical Customer (m^3)	Annual Bill for Typical Customer[1] ($)	DSM Amount for Recovery[2] ($)	Estimated % of Annual Bill
Rate 1 - Heating & Water Heating	2,400	$824	$6	0.7%
Rate 6 - Commercial, Heating & Other Uses	22,606	$5,996	$2	0.0%
Rate 9 - Container Service[3,5]			$331	
Rate 100 - Commercial, small size	339,188	$71,270	$0	0.0%
Rate 110 - Industrial, small size, 50% Load Factor	598,568	$112,916	$539	0.5%
Rate 110 - Industrial, avg. size, 75% Load Factor	9,976,120	$1,720,770	$8,977	0.5%
Rate 115 - Industrial, small size, 80% Load Factor	4,471,609	$752,389	$0	0.0%
Rate 125 - Extra Large Firm Distribution[4,5]			$2,475	
Rate 135 - Industrial, Seasonal firm	598,567	$97,464	($1,143)	-1.2%
Rate 145 - Commercial, avg. size	598,568	$107,780	($7,987)	-8.0%
Rate 170 - Industrial, avg. size, 75% LF	9,976,120	$1,490,389	($31,156)	-2.1%
Rate 200 - Wholesale Service[3,5]			$4,291	
Rate 300 - Firm or Interruptible Distribution[4,5]			$412	

1. Annual bills based on October 1, 2017 rates.

2. DSM amounts for Recovery do not include interest amounts that will apply at the time of clearing.

3. Information is for the total amount for DSM recovery.

4. DSM amounts for recovery for Rate 125 and Rate 300 are for average customers in each rate class.

5. Rates 9, 125, 200 & 300 will not have any LRAM component included in the rate allocation since customers in these rates classes are not eligible for DSM programs. These rate classes will however, be subject to rate allocations for DSMVA and applicable DSMIDA related to the Low Income Program.

ENBRIDGE RESPONSES TO 2015 ANNUAL VERIFICATION RECOMMENDATIONS

The following is a summary of the recommendations provided by the Evaluation Contractor in the 2015 Annual Verification and Enbridge's responses where applicable to Enbridge.

Table 5-1 Summary of recommendations that apply to the overall annual verification[1]

#	Overall Annual Verification Recommendation	Applies to		
		Union	Enbridge	Evaluation
O1A	Consider investing in a relational program tracking database.	✓	✓	
O1B	Enbridge should include site-level information for all measures installed through the program.		✓	
O2A	Deliver tracking data in a single flat file.	✓	✓	
O2B	Consider investing in a relational program tracking database.	✓	✓	
O3A	Develop and maintain an electronic summary of the TRM.	✓	✓	✓
O3B	Track prescriptive savings using unique measure descriptions that map to electronic TRM.	✓	✓	✓

[1] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

O1. Finding: The Enbridge tracking database does not currently include information that allows the evaluator to identify all the projects installed by a single customer. Without this information, the EC could not identify projects installed across customers to determine whether interactive effects may have reduced energy savings. Some prescriptive measures in the Enbridge data did not have site-level information at all, only a summary of the energy savings for that technology across all sites.

Recommendation A: Both utilities should strongly consider investing in relational program tracking databases. Relational program tracking databases and customer relationship management ("CRM") systems allow for multiple measures and projects to be associated with a single customer and/or customer site. The incremental cost of implementation is low if it is part of the initial database design, populated as projects are started, and updated once they are complete.

Outcome: Reduced burden on utility staff and reduced evaluation costs. A relational database would streamline aggregation of program data for scorecards and make providing data simpler for annual savings evaluation and verification.

ENBRIDGE RESPONSE: As detailed in its 2015 to 2020 Multi-Year Plan, Enbridge outlined the need for a DSM IT system replacement. The Board approved this request in its January 20th, 2016 Decision. As a result, Enbridge DSM is currently undergoing a system upgrade that will include improved tracking & reporting and CRM components. This system upgrade is expected to be rolled out in 2018.

Recommendation B: Enbridge should include site-level information for every measure installed in the program.

Outcome: Confirmation that each installation is unique.

ENBRIDGE RESPONSE: Though the summary tracking information initially provided to the EC for quasi-prescriptive measures, in some cases, did not include all site-level information, upon request Enbridge provided the EC with all the detailed information maintained in back up documentation for each project. Enbridge will endeavour to include comprehensive information for every measure in its summary tracking data moving forward. Given the timing of the receipt of the 2015 Annual Verification Recommendations, in Q3 of 2017, after the completion of the 2016 program year, efforts to make significant changes to tracking for the already completed 2016 program year will be limited, however Enbridge will work to ensure all of the information requested is included in the tracking data summarized to the EC.

O2. Finding: Both utilities invested significant effort in developing Excel-based tracking workbooks that summarized data and calculated DSMSI based on utility-reported results. Union's workbook included a feature that was designed to allow evaluators to enter adjustment factors in a single location and automatically update DSMSI and LRAM calculations. Neither workbook was well suited for evaluation efforts.

Recommendation A: Deliver to evaluators a single, flat file of tracking data.[2] Each record should have measure-level information which includes the information listed below.

- Program identification information, such as scorecard, and program name
- Customer identification information, such as a unique customer ID, rate class, and location
- Measure identification information, such as measure description, unique

[2] In this context, a flat file is a table with one record per line and no summary information

measure identification, measure group, measure life, free rider rate, and savings per unit for prescriptive measures

- Savings information, such as annual gross and net savings, cumulative gross and net savings, and non-gas savings
- Additional information as needed to allow the evaluator to verify LRAM and cost-effectiveness

The Union tracking data most closely followed this recommendation, but both utilities invested in workbook features that did not enhance evaluation efficiency.

Outcome: Reduced burden on program staff, more flexibility for evaluators.

ENBRIDGE RESPONSE: With the exception of some quasi-prescriptive measures, project related measure-level information was included in the original tracking database provided by Enbridge to the EC for the 2015 Verification (all requested information was ultimately provided to the EC). Consistent with Enbridge's presentation of results during the 2012-2014 DSM Framework, the Enbridge tracking and reporting summary provided to the EC included dynamic calculation tools that linked measure level inputs to the energy savings calculations, cost-effectiveness calculations, scorecard achievements, and shareholder incentive calculations, as well as LRAM impacts for the 2015 program year. In line with the EC comment in this finding, like Union's workbook, Enbridge's workbook included a feature that was designed to allow evaluators to enter adjustment factors in a single location and automatically update DSMSI and LRAM calculations, Enbridge's tracking summary has evolved and improved through the review of previous audits to a comprehensive and transparent tool. Prior auditors and Audit Committees expected Enbridge's tracking database to have this level of transparency to fully illustrate the determination of scorecard achievements. Given the timing of the receipt of the 2015 Annual Verification Recommendations, after the completion of the 2016 program year, efforts to make significant changes to the tracking tool for the already completed 2016 program year are challenging, however Enbridge is making every effort to ensure the 2016 tracking summary clearly provides the information requested.

Recommendation B: See recommendation O1A. The utilities should consider investing in a new database.

Outcome: Reduced burden on utility staff and reduced evaluation costs.

ENBRIDGE RESPONSE: See Enbridge Response O1A.

O3. Finding: Neither Union nor Enbridge tracking databases currently use prescriptive measure descriptions that map directly to the approved energy savings spreadsheet ("TRM"). The EC often struggled to align tracking measures to the correct TRM measure, which resulted in repeated back-and- forth between evaluation and the utilities for clarification. During this process, the EC found that some Enbridge measures were assigned to the wrong sub-category by capacity or other size measure. The EC also found that some Enbridge measures were assigned outdated savings values from previously- approved TRMs.

Recommendation A: Develop and maintain an electronic summary of the TRM, such as an Excel file. Each measure (identified as a unique savings value) should have an assigned measure ID number, and new ID numbers should be assigned when a measure is updated with a new savings value. This allows for a historical record of the changes in the TRM and allows the evaluation to identify outdated values.

ENBRIDGE RESPONSE: Board staff now coordinates the TRM update process. [3]

Recommendation B: Track prescriptive savings using unique measure descriptions that clearly map to the electronic TRM.

Outcome: Reduced burden on utility staff and reduced evaluation costs. Fewer errors in the tracking data.

ENBRIDGE RESPONSE: Enbridge will work to provide the EC with a clearer mapping of prescriptive measure descriptions in its tracking database to measure descriptions outlined in the TRM.

[3] As outlined in the Board's March 4 letter, regarding the Transition of Technical Evaluation Committee Activities to the OEB the online portion of the TRM has been transitioned to OEB staff.

Table 5-2 Summary of recommendations that apply to RunitRight[4]

#	RunitRight Recommendation	Applies to		
		Union	Enbridge	Evaluation
RR1	Consider adding independent variables to the regression to account for school breaks.		✓	
RR2A	Consider including the date when each activity was implemented.		✓	
RR2B	Provide information on both the baseline and installed case.		✓	
RR2C	Increase the level of documentation when a single change results in a significant portion of savings.		✓	
RR3A	Consider including a basic description of all end-use equipment served by the gas meter.		✓	
RR3B	Consider using engineering calculations to estimate electricity savings.		✓	
RR3C	Consider reviewing the process for selecting the HDD reference temperature.		✓	

RR1. Finding: Not all the RunitRight regression models provided a strong fit for the consumption data. In particular, school buildings, which have widely inconsistent occupancy throughout the year, show low R- squared values.

Recommendation: Consider including additional independent variables for schools to account for break periods, which may improve the regression fit.

[4] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

Outcome: More confidence in the reported savings estimates.

ENBRIDGE RESPONSE: Where the consideration of baseline period is required to facilitate a regression model, to address the unique occupancy of school buildings, moving forward the selected baseline period for school projects is September to August and adjusted as appropriate depending on the date of implementation and data availability.

RR2. Finding: The RunitRight documentation includes a description of the activities at each site, which are documented in the calculation workbook and annual site report. The same level of documentation is included for all activities, regardless of the percentage of savings contributed by that activity.

Recommendation A: Consider including the date when each activity was implemented.

ENBRIDGE RESPONSE: Given the timing of the receipt of the 2015 Annual Verification Recommendations, in Q3 of 2017, beginning with the 2018 program year Enbridge will work to include the implementation date for each activity in the project documentation.

Recommendation B: Provide information on both the baseline and installed case. For example, when a schedule is reset, provide the pre- and post-installation schedule.

ENBRIDGE RESPONSE: Given the timing of the receipt of the 2015 Annual Verification Recommendations, in Q3 of 2017, beginning with the 2018 program year Enbridge will look at how to supplement the project file to include additional information and details on both the baseline and installed case.

Recommendation C: Increase the level of documentation on end use equipment when a change to that equipment results in a significant reduction in consumption.

ENBRIDGE RESPONSE: The Run it Right offer is focused on achieving gas savings through the optimization of existing building systems and equipment through the implementation of low cost/no cost improvements to a building's operation as identified through the offer's investigation process and monitored through the support of an EMIS. Gas savings are determined based on a holistic comparison, where savings are determined through a regression analysis of the consumption data impacted by the combination of all measures implemented and not attributed to any specific measure.

Outcome: More confidence in the reported savings estimates.

RR3. Finding: The evaluator observed a number of opportunities to improve the savings estimates associated with the RunitRight program, including savings at the electric meter. Some sites had base loads that were unexpectedly sensitive to the reference temperature.

Recommendation A: Consider including a basic description of the end-use equipment served by the gas meter, such as DHW, heating, or cooking. This will help the reviewer better assess the consumption patterns occurring over time and the magnitude of base load and weather-sensitive savings estimated.

ENBRIDGE RESPONSE: Given the timing of the receipt of the 2015 Annual Verification Recommendations, in Q3 of 2017, Enbridge will work to incorporate this recommendation beginning in 2018.

Recommendation B: Consider using engineering calculations to estimate electric energy savings to capture the full value of the program.

ENBRIDGE RESPONSE: Though Enbridge recognizes that capturing electric savings would demonstrate additional value from the offer, the Run it Right offer will continue to focus on the determination of low cost/no cost gas savings that are identified through building optimization recommendations.

Recommendation C: Consider reviewing the process for selecting the HDD reference temperature to reduce baseload sensitivity.

Outcome: More accurate savings estimates.

ENBRIDGE RESPONSE: Given the timing of the receipt of the 2015 Annual Verification Recommendations, in Q3 of 2017, Enbridge will investigate the process for selecting the HDD reference temperature beginning in 2018.

Table 5-3 Summary of recommendations that apply to simulation modeling[5]

#	Simulation Modeling Recommendation	Applies to		
		Union	Enbridge	Evaluation
SM1	Provide simulation file and output to the evaluation team.	✓	✓	
SM2	Provide more explicit support for major measure installations.	✓	✓	
SM3	Consider reviewing and modifying program processes to avoid data entry or outdated simulation result errors.	✓	✓	
SM4	Consider funding a study to verify the models produced by the utility agents.			✓

SM1. Finding: Both utilities use building simulation modeling to estimate energy savings for their home retrofit programs, including Home Energy Conservation, Home Reno Rebate, Winterproofing, and the Home Weatherization Program. HOT2000 is the most common program used for those simulations, which is a program developed and released by NRCan for certified energy advisors. Because of the restrictions on the program, the evaluator could not consistently run the simulation files and produce the same result reported by the program.

Recommendation: Provide both the building simulation file and the program output to the evaluation team. By delivering both, the evaluation team would not have to follow

[5] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

up with the utility to obtain output for models that could not be run, but could still verify the output for models that can be run.

Outcome: Reduced burden on utility staff and reduced evaluation costs.

ENBRIDGE RESPONSE: In the 2015 Verification, Enbridge provided the HOT2000 files to the EC as per the EC's request. When the EC subsequently indicated to Enbridge that they were experiencing some difficulty opening these files, Enbridge offered to provide excel files which provided an extract of data contained in the EnerGuide HOT2000 files. Enbridge would be pleased to provide both the HOT2000 files and the excel file containing the HOT2000 file outputs in any future request.

SM2. Finding: Both utilities have market-rate scorecard metrics that rely on a definition of deep savings that is related to the number of "major" measures installed at a site. Both utilities also collect and deliver photographs to support many of the changes made at a home retrofit site. However, the evaluator could not consistently confirm the number or type of major measures installed based on the photographs or other documentation provided.

Recommendation: Consider providing more explicit support for each major measure to eliminate uncertainty around the number of deep savings program participants.

Outcome: Greater certainty around scorecard achievements.

ENBRIDGE RESPONSE: Enbridge endeavours to provide all available supporting information to the EC as requested. The Home Energy Conservation ("HEC") offer is modelled after NRCan's Residential ecoENERGY Retrofit program which ran until March 2012. Supporting information gathered for measures installed through the

HEC offer is consistent with what NRCan continues to require of all certified Energy Advisors for use of NRCan's licensed HOT2000 energy modelling software in EnerGuide mode, for example invoices or receipts to support upgrades and supporting photographs. Enbridge commits to continue to work to ensure the EC has the available information to facilitate the confirmation of the number or type of major measures installed in a home undergoing review.

SM3. Finding: The evaluator identified a number of inaccurate savings entries due to data entry errors or outdated Union home retrofit simulation results. Many of these errors could be avoided through changes in program processes.

Recommendation: Consider reviewing and modifying program processes to avoid similar errors in the future.

Outcome: Reduced burden on utility staff and reduced evaluation costs.

<u>ENBRIDGE RESPONSE</u>: Enbridge understands it had minimal inaccurate savings entries due to data entry errors or outdated simulation results, however in line with the utility's objective of continuous improvement, Enbridge will carry on working to increase accuracy wherever possible. Notwithstanding the EC's recommendations, in each of the 2016 and 2017 program years, Enbridge has expanded tracking and reporting including a deeper analysis of EnerGuide data exports to identify data entry errors,

SM4. Finding: The energy savings from the home retrofit programs rely exclusively on the simulations provided by the delivery agents. Those simulations likely rely on a number of assumptions or standard modeling practices which may or may not follow industry standards.

A detailed review of the models was outside the scope of the annual audit.

Recommendation: Consider funding a study to verify the models produced by the utility agents to ensure they conform to standard industry practice.

Outcome: Greater certainty around savings estimates.

ENBRIDGE RESPONSE: This recommendation was not directed to Enbridge but for clarity, HOT2000 is developed and managed by the Office of Energy Efficiency at Natural Resources Canada ("NRCan"). HOT2000 simulations in EnerGuide mode can only be completed by Energy Advisors who have been certified by NRCan to use NRCan's HOT2000 modelling software. Periodically, NRCan updates the software to reflect learnings and implement improvements. Most recently in 2017, NRCan released the newest version, HOT2000 V. 11.3 and EnerGuide Rating System ("ERS") V. 15.1. In order to deliver services and perform energy audits using this version, NRCan expects energy advisor candidates to demonstrate proficiency by passing the Foundation Level exam; passing the Energy Advisor exam, and be affiliated with a service organization; and they must complete probationary HOT2000 files to the satisfaction of the service organization to show competence with energy simulation modeling and field work in addition to training on current NRCan industry standard inputs and modeling practices. All HOT2000 simulation files, once completed by certified energy advisors for HEC, are provided by the service organizations to NRCan and are subject to NRCan's QA procedure.

Table 5-4 Summary of recommendations that apply to cost-effectiveness[6]

#	Cost-effectiveness Recommendation	Applies to		
		Union	Enbridge	Evaluation
CE1	Allocate "sector"-level administrative costs and overhead to each individual program and report program-level cost-effectiveness results.	✓	✓	
CE2	Use a consistent real discount rate of 4% when using real streams of benefits and costs.	✓	✓	
CE3	Explore the possibility of better defining water avoided costs.	✓	✓	✓
CE4	Work towards better uniformity in methods and assumptions.	✓	✓	✓

CE1. Finding: In some cases, the Union program costs were grouped together for several programs. To get program- or sector-level cost-effectiveness results, the EC prorated costs to programs based on natural gas savings.

Recommendation: Allocate "sector"-level administrative costs and overhead to each individual program and report program-level cost-effectiveness results.

Outcome: Greater certainty around program-level achievements.

<u>ENBRIDGE RESPONSE</u>: Enbridge will continue to work to appropriately allocate DSM costs, practically and reasonably, in line with direction provided in the Board's

[6] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

Guidelines, including for the purpose of conducting cost-effectiveness screening. As outlined in section 9.1.2 of the Guidelines:

> For the purpose of the TRC-Plus test, the Program Costs relate [sic] to DSM program [sic] include the following components:
> i) Development and Start-up;
> ii) Promotion;
> iii) Delivery;
> iv) Evaluation, Measurement and Verification ("EM&V") and Monitoring; and
> v) Administration.
> Of the above costs, only Start-up, Promotion, Delivery, some Evaluation and Verification are applicable to individual programs. Other costs related to the design and delivery of DSM programs, are appropriately considered at the DSM portfolio level. These include Development, some Evaluation costs, and Monitoring, Tracking and Administration costs.[7]

The Guidelines further specify, "for practical purposes, if certain administrative costs cannot be assigned to individual programs these costs should be accounted at the portfolio level."[8]

CE2. Finding: Enbridge uses a real discount rate of 4% and applies it to streams of current (nominal) values. However, the real discount rate should only be applied to real (inflation-adjusted) streams of benefits and costs. Nominal discount rates should be applied to streams of current (nominal) values.

Recommendation: Use a consistent real discount rate of 4% for both Enbridge and Union when using "real" (inflation-adjusted) streams of benefits and costs.

Outcome: More accurate cost-effectiveness results.

---

[7] Filing Guidelines to the 2015-2020 DSM Framework, EB-2014-0134, page 28.
[8] Ibid. page 29

<u>ENBRIDGE RESPONSE</u>: Enbridge agrees it is appropriate to apply the real discount rate of 4% to real (inflation-adjusted) streams of benefits and costs. Enbridge further agrees it is appropriate to apply nominal discount rates to streams of current (nominal) values.

CE3. Finding: Water rates are currently used as a proxy for the water avoided costs. Water avoided costs should only include the marginal impact from reduced consumption. Using the full rate as the avoided cost may be appropriate in some jurisdictions with a completely variable rate structure. However, those with high fixed costs (which, in our experience, can represent 75% to 80% of water costs) should use a true avoided cost.

Recommendation: Explore the possibility of better defining water avoided costs.

Outcome: More accurate cost-effectiveness results.

<u>ENBRIDGE RESPONSE</u>: Enbridge agrees that water avoided costs should include only the marginal impact from reduced consumption. As part of the 2015 verification, the EC recommended an adjustment to Enbridge's avoided water costs to reflect a more appropriate estimate of avoided water costs. Enbridge applied a similar approach with its 2016 avoided water costs.

CE4. Finding: The EC found major discrepancies in the way the utilities calculate cost-effectiveness. Some areas of discrepancies included the discount rate, the use of a non-energy benefit adder, the format of reporting results, and the allocation of administration and overhead costs by program. While there is always a balance to be found between uniform methods and the need to account for each specific utility's needs, greater uniformity could be achieved.

Recommendation: Work towards a better uniformity of cost-effectiveness methods and assumptions between the two gas utilities.

Outcome: More accurate and consistent cost-effectiveness results.

ENBRIDGE RESPONSE: Enbridge acknowledges the EC's recommendation and moving forward, Enbridge will consult with Union in an effort to work towards better uniformity of cost-effectiveness methods and assumptions.

Table 5-5 Summary of recommendations that apply to other areas[9]

#	Other Recommendation	Applies to: Union	Applies to: Enbridge	Applies to: Evaluation
OR1	When the C&I deep savings metric is used, deliver monthly billing data for each C&I participant.	✓		
OR2	Provide a detailed explanation for the DSMSI calculation.	✓	✓	

OR1. Finding: The Union scorecard includes a metric that relies on an understanding of the whole- building energy use for each C&I program participant. The program data included the total annual consumption at each site, normalized by a regional (north or south) estimate of heating degree days. The calculation appeared to assume that industrial sites were not weather-sensitive but commercial sites were.

[9] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

Recommendation: When the C&I deep savings metric is used, deliver monthly billing data for each C&I participant to allow the EC to verify the annual consumption values and the weather sensitivity assumptions. Provide the supporting information (and calculation, if possible) for the normalized regional heating degree days.

Outcome: Greater certainty around scorecard achievements.

ENBRIDGE RESPONSE: Not Applicable

OR2. Finding: The evaluator was unable to locate a source document that supports the utilities' calculation of DSMSI. Given the importance of the shareholder incentive, it is appropriate to have a clearly defined and detailed explanation of how it is calculated.

Recommendation: Provide a detailed explanation for the DSMSI calculation for review by the EC and OEB.

Outcome: Greater certainty around shareholder incentives.

ENBRIDGE RESPONSE: This recommendation was previously addressed in the course of the 2015 verification process. Enbridge provided a detailed explanation regarding the calculation of the shareholder incentive to the EC and the EAC during the course of the 2015 verification. The approach followed the calculation outlined by the Board in the previous Guidelines, EB-2008-0346. The calculation to be used by Enbridge in 2016 was subsequently provided and has also been reviewed with the EC and EAC.

Table 5-6 Energy savings and program performance recommendations[10]

#	Energy Savings and Program Performance Recommendation	Applies to		
		Union	Enbridge	Evaluation
ES1	The utilities should continue in their commitment to accuracy.	✓	✓	
ES2	Evaluate free-ridership for the programs annually and couple the free-ridership evaluation with process evaluation			✓
ES3	Error ratios from this report inform sample design for future evaluation.			✓
ES4	Align the program design with cumulative net goals	✓	✓	
ES5	Do not pay incentives until after installation is complete.	✓	✓	
ES6	Develop policies to collaborate across electric and gas projects to avoid double-counting fuel savings and increases from energy efficiency measures.	✓	✓	
ES7	Consider establishing a policy to define rules around energy savings calculation for fuel switching and district heating/cooling measures.	✓	✓	✓
ES8	Consider establishing a policy that defines an eligibility floor and cap based on simple payback period for energy efficiency projects.	✓	✓	
ES9	Consider establishing an official definition for EUL and implementing a study to define EULs for program measures	✓	✓	✓
ES10	Track metrics for how long it takes from the final installation verification to the posting of incentive payments.	✓	✓	
ES11	Increase transparency of "influence adjustments" and do not include in gross	✓		

10 EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

#	Energy Savings and Program Performance Recommendation	Applies to: Union	Enbridge	Evaluation
	savings			
ES12	Conduct a process evaluation to improve Large Volume influence on customer projects	✓		
ES13	Consider approaches to market that leverage third-party vendors.	✓	✓	

ES1. Finding: Both utilities exhibit a strong commitment to accurate energy savings estimates. Both utilities have made significant investments in developing calculation tools which model savings accurately. For example, Union's dock door seal calculator is well considered and designed, and Enbridge's Etools calculator is very thorough in attempting to model savings for key measures. Both utilities chose to retain engineers with strong understandings of their customers' building and process systems. We had numerous opportunities to interact with these engineers on phone calls and site visits, and have grown to respect their knowledge and engagement with the types of systems that matter to their customers. Both utilities showed a commitment to finding accurate savings. On several occasions, both on the phone and in writing, the evaluation team suggested a value that would have increased savings in a way that the program engineer did not think was valid. When this happened, neither utility was shy in suggesting that we may want to make a more conservative choice.

Recommendation: The utilities should continue in their commitment to accuracy.

Outcome: Accurate energy savings.

ENBRIDGE RESPONSE: Enbridge is committed to continue striving for accurate savings calculation estimates. Enbridge has been a leader in refining savings calculations for many technologies and will continue to look for opportunities to improve approaches and calculation tools with consideration for new information and learnings.

ES2. Finding: Free-ridership in the utilities' programs is high

Recommendation: With high free-ridership and rapidly changing programs, consistent evaluation of free-ridership annually and free-ridership evaluation coupled with process evaluation will help identify specific ways for each program to manage and reduce free-ridership.

Outcome: Effective free-ridership management will allow the programs to increase their net savings significantly in future years.

ENBRIDGE RESPONSE: This recommendation was not directed to Enbridge however, Enbridge finds it necessary to make clear that it does not have confidence in the ECs findings. There are well documented failings and concerns with self-report survey approaches that were proven out in the EC's findings. Enbridge notes there were significant concerns with the measurement of free-riders conducted by the EC on the 2015 program year custom offers. The study did not, in a number of instances, reflect industry best practice. Also, given the limited information shared with the utility with respect to how survey responses were interpreted and translated into scores, and with no way to know if calculations of NTG scores were done corrected, Enbridge does not have confidence in the results. Enbridge is particularly concerned with the reliability of scoring that was determined based on feedback from customer representatives regarding projects undertaken up to 2 ½ years earlier. Beyond the ability to reach truly informed participants to the projects, Enbridge is concerned the delayed research significantly exacerbated the inherent recall bias of survey participants. In addition, the limited research conducted to ascertain utility influence delivered through the utility's business partners; the focus in the survey instrument on the payment of incentives rather than assessment of the entire suite of services and values provided to customers; the posing of questions that did not ensure clarity of properly captured efficiency improvements relative to specific project utilized baselines to ensure there could be no double counting of adjustments; and, the dismissal of consideration of utility influence and long standing customer support prior to the current program year, collectively contributed to proposed free-rider values in which there can be no confidence.

Notwithstanding Enbridge's concerns with the NTG study findings and the approach taken in the 2015 evaluation effort, given the EC's recommendation here, and the considerable discussion on this process during the 2015 verification, it is puzzling that OEB Staff has decided to not proceed with planned free-ridership/NTG evaluation on custom programs in the following year's evaluation.

ES3. **Finding:** Relative precision targets were exceeded for some programs and not met for others.

Recommendation: Error ratios from the results provided in this report should be used to inform sample design for future evaluation years.

Outcome: Better defined error ratios for the measures in the programs will allow more efficient sample design for future evaluations, improving precisions and reducing costs.

<u>ENBRIDGE RESPONSE</u>: This recommendation was not directed to Enbridge however, Enbridge has shared its concerns with the EC and the EAC regarding the error ratios in the results presented in the CPSV and NTG sampling and the resulting uncertainties regarding any accuracy in the proposed adjustments.

ES4. Finding: Attribution for the programs came primarily through acceleration rather than changes in efficiency or quantity/size. This is partly due to the measures that dominate the programs: controls, maintenance, and optimisation. These measures do not have varying efficiencies, so the programs are either affecting the number of units implemented or accelerating the measure. Acceleration is less valuable to programs that are seeking to meet cumulative net goals. Acceleration periods tend to be considerably shorter than the estimated useful life ("EUL") of a measure and thus the partial attribution that results is low relative to cumulative gross savings.

Recommendation: To align the programs with cumulative net goals, the utilities should seek to:

- continue promoting long life measures and consider discontinuing promotion of short lived measures
- proactively upsell equipment purchases from standard to efficient products
- target hard to reach customers who have not participated in the past
- promote EE measures with low market penetration (such as heat reflector panels)
- motivate customers to increase the scope of their projects, some options include multi-measure bonuses or escalating incentive structures that pay more

for doing more

Outcome 1: Focusing on proactive sales rather than reactive will help the net-to-gross ("NTG") ratio.

Outcome 2: Effective free-ridership management will allow the program to increase net savings significantly in future years.

ENBRIDGE RESPONSE: While Enbridge does focus its efforts on achieving cumulative gas savings, and agrees to continue to work to align programs to seek improvements in areas recommended above by the EC, Enbridge also intends to continue to deliver programming to support a wide range of eligible energy conservation projects to address the multiple key priorities set out by the Board. The Framework specifically stated that:

> DSM budgets will be driven by the gas utilities' ability to increase activity and address the key priorities discussed below, including delivering tailored service to those customers who have already increased their efficiency levels but can continue to realize savings, increasing operational efficiency improvements, and incorporating behavioural changes into program offerings.[11]

ES5. Finding: A handful (<5) of respondents indicated that all or part of their incentivized project had not yet been installed over a year after the incentive was paid.

Recommendation: Do not pay incentives until after installation is complete.

Outcome: Cost-effectiveness of the program will increase as it avoids paying for savings that do not materialize.

---

[11] EB-2014-0134, Report of the Board, DSM Framework for Natural Gas Distributors (2015-2020), Section 4.2, page 19

<u>ENBRIDGE RESPONSE</u>: Enbridge did not receive details from the EC identifying any projects where this was the case. Enbridge is unaware of any specific project(s) that received incentive payments having not yet been installed. Enbridge requires that projects are completed prior to the payment of incentives.

ES6. Finding: Some customers receive incentives from their electric provider and natural gas utility to complete the same EE measure. Both providers may claim the same changes in energy use, resulting in overlap when aggregated across fuels at the provincial level.

Recommendation: Develop policies to collaborate across electric and gas projects to avoid double- counting fuel savings and increases from energy efficiency measures.

Outcome: More accurate energy and carbon savings estimates across the province.

<u>ENBRIDGE RESPONSE</u>: Enbridge is not aware of the EC providing details regarding any observations of double counting of fuel savings for custom projects that were reviewed through the course of the 2015 Verification. As outlined in the Board's Framework and Guidelines (EB-2014-0134), Enbridge continues with efforts to co-ordinate DSM and CDM programs and increase collaboration with electricity programs where possible and appropriate.

ES7. Finding: Some measures (e.g., geothermal heat pumps, combined heat and power, and those that save district heating energy) have difficult-to-define baseline technologies.

Recommendation: Consider establishing a policy to define rules around energy savings calculation for fuel switching and district heating/cooling measures.

Outcome: Less evaluation risk and a better alignment between province energy efficiency goals and program implementation.

ENBRIDGE RESPONSE: Enbridge will look at considerations to define approaches to energy savings calculations for fuel switching and district heating/cooling measures.

ES8. **Finding:** Projects with very long and very short simple payback periods often have low NTG ratios. However, from a customer service standpoint, it may be difficult for utilities to deny incentives to customers unless they have pre-established rules to point to.

Recommendation: Consider establishing a policy that defines an eligibility floor and cap based on simple payback period for energy efficiency projects.

Outcome: The rule will give utilities a guideline to restrict the program to projects that are more likely to result in net savings. It will also allow the utilities to reject potentially poor projects without a large effect on customer satisfaction.

ENBRIDGE RESPONSE: As set out in the Board's Decision and Order, Section 5.2.6 on Enbridge's 2015-2020 Plan (EB-2015-0049), the OEB rejected the need to introduce a policy defining payback eligibility criteria for the Commercial and Industrial custom offer.

ES9. **Finding:** Members of the EAC and evaluation team have different understandings of the definition of some evaluation inputs.

Recommendation: Consider establishing an official definition for EUL and implementing a study to define EUL for all measures, especially steam traps, pipe leaks, steam leaks, condensate leaks, and pipe insulation.

Outcome: The study will improve the accuracy of lifetime savings estimates.

ENBRIDGE RESPONSE: A definition for EUL was included in the glossary of terms developed as part of the Board approved TRM filed in December 2016. Enbridge understands that Board Staff has issued an RFP for a CI custom measure life review as part of the 2016 evaluation process.

ES10. Finding: A handful (<5) of sites reported unhappiness with delays in receiving their incentive payment (5 months).

Recommendation: Track metrics for how long it takes from the final installation verification to the posting of incentive payments. Consider holding program managers accountable to these metrics by considering them during performance reviews, building in performance bonuses if all payments are posted within one month, and/or implementing a penalty if it takes greater than three months to post any payments.

Outcome: Improved customer satisfaction.

ENBRIDGE RESPONSE: Enbridge requires that projects must be completed prior to the payment of incentives. Incentives are paid only after the measure(s) are installed, and the project is completed and fully commissioned. Enbridge is unaware of any customer complaints regarding payment delays. In any case where such an observation has been made, Enbridge suggests a review of the specific circumstances is in order to confirm that the customer had completed and submitted all project requirements necessary to meet project completion standards and facilitate timely payment; this may help clarify the circumstances for any identified delays.

ES11. Finding: Influence adjustments were made to projects that adjusted the gross savings for “net” or program influence reasons. Accounting of which projects had these adjustments was not maintained by the program and the adjustments were included in different places in project calculation workbooks, making their identification challenging. In addition, the program NTG was also applied to these projects, effectively double discounting savings in scorecards.

Recommendation: If the utility chooses to continue making influence adjustments to the savings upon which it calculates savings, these adjustments should be made more transparent and not included in the reported gross savings for the program in scorecards. Instead the specific project influence adjustment should be included in the scorecard in place of the general program or domain level NTG factor.
Outcome: Reduced risk of double adjustments.

ENBRIDGE RESPONSE: Not Applicable

ES12. Finding: Union’s Large Volume program has a very high amount of free-ridership.

Recommendation: This evaluation did not include a process evaluation. Union should consider conducting a process evaluation focused on how to reduce the rate of free-ridership. Three options that the Union might consider are:

- Eliminate measure types with high free-ridership (Union indicated that most maintenance type measures were eliminated in 2016).
- Use an application process that includes a committee review that can reject free riders. This option is hard for utilities to manage as it can affect customer satisfaction negatively
- Clear payback criteria such as initial payback must be longer that X years and the incentive paid must reduce payback below Y years. This has the advantage of being a rule that account representatives can explain when talking to customers.
- Non-energy benefits of projects that large industrial customers gravitate to are often large compared to energy saving benefits, so simple payback criteria will not eliminate all free rider projects. Awareness of this issue should be promoted among the implementation team.

Outcome: Effective free-ridership management may allow the program to increase its net savings significantly in future years.

ENBRIDGE RESPONSE: Not Applicable

ES13. Finding: Vendor attribution did not increase overall program attribution significantly. Of the vendors that customers cited as influences, few indicated that either program had much effect on the projects.

Recommendation: The utilities should consider approaches to market that leverage third-party vendors. A process evaluation that includes vendor interviews might uncover opportunities.

Outcome: Effective leveraging of vendors could both increase NTG ratios and increase program uptake.

ENBRIDGE RESPONSE: Enbridge's approach to market for its commercial and industrial offers fully leverages third party vendors. For many years, Enbridge has extensively engaged business partners including vendors/contractors/engineers and distributors to promote the Enbridge DSM program and support customers in the decision making process, propelling customers to implement energy efficiency improvements. Enbridge has found this approach to be highly effective in extending the utilities reach and increasing project uptake.

Table 5-7: Verification process recommendations[12]

#	Verification Process Recommendation	Applies to		
		Union	Enbridge	Evaluation
VP1	Modify contracts to require participants to agree to comply with EM&V as well as utility representatives as part of the requirements for participation in the program.	✓	✓	
VP2	The verification and utility staff should agree to a code of conduct for each role during onsite visits.	✓	✓	✓

VP1. Finding: DNV GL was unable to obtain access to all the equipment at all the sites selected for verification. Both Enbridge and Union have several large projects with industrial companies, including food processing, refineries, and other industries. In

[12] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

many cases, the customer refused to provide SCADA data or similar trend data to allow a reasonable verification of the project. This means we were unable to do more than a reasonableness check on the savings**.**

A review of the Enbridge contract shows that the customer is not required to provide the information that is necessary for EM&V. The most relevant sections are:

- Item 6 states: Payment of the Incentive Payment is subject to the completion of a satisfactory site inspection of the improvements, including the installed equipment by an authorized representative of Enbridge.
- Item 9 states: Upon request within eighteen months of the commissioning date of the Project, and with reasonable notice, the Customer agrees to provide authorized representatives of Enbridge with access to the Project, and with required information or data relating to the project for the purposes of the Application and these General Terms and Conditions.

Neither of these are sufficient for EM&V.

Recommendation: Modify contracts to require participants to agree to comply with EM&V as well as utility representatives as part of the requirements for participation in the program.

Outcome: Reduced evaluation costs and risks. Participant non-compliance requires evaluators to request documentation for a large backup sample, and to survey and/or visit additional sites to obtain sufficient data for the evaluation. The process of contacting a site and getting a refusal costs time and money, as does the substitution of an additional site to make up for the unobtained data. In some cases, there might not be additional sites to sample, in which case the evaluation estimates will have lower precision than they would with full compliance.

<u>ENBRIDGE RESPONSE</u>: Enbridge encourages its customers to comply, cooperate and participate with all EM&V verification activities. At the same Enbridge recognizes it is important to be respectful that customers are busy running business and requests for customers' time should not be overly burdensome. Up until this 2015 Verification, virtually 100% of sampled participants selected for verification have complied with verification related requests. In the 2015 sample, in some cases, Enbridge received feedback from customers that onerous time requirements and/or specific data requests made of customers may not have been considered reasonable and/or comprised customer privacy concerns or safety policies. In addition, the delay between project completion and third party evaluation, of greater than 2 years in some cases, may have further discouraged customers to participate fully in the 2015 Verification because the appropriate person that should respond was now not available. Enbridge believes the language contained in Item 9 in Enbridge's Energy Efficiency Project Application General Terms and Conditions details that the customer has agreed to allow access to the project and the required information or data relating to the project as a condition of participation. Enbridge will investigate however how it might improve the language.

VP2. Finding: Verification engineers and verification forms caused confusion with site contacts and the length of visits also led to a handful of customer complaints. Utility staff at a handful of sites responded to questions in place of participating customers and in one case interfered with data collection.

Recommendation: The verification and utility staff should agree to a code of conduct for each role. The teams should receive clear direction as to the dos and don'ts of all parties involved in site visits, including both verification engineers and utility staff should they attend the visit. Open lines of communication between the site team and utility staff should be maintained to reduce misunderstandings and ensure that the teams are on the same page as to each other's role.

In general, the following should be part of standard verification practices:

- Ensure site engineer reviews final site report for accuracy post-audit.
- Align data collection forms with site report structure to reduce communication and transcription errors.
- Ensure data appropriate to determining EUL is collected while on-site (i.e., make EUL determination a primary, rather than secondary focus).
- Request specific documentation or data from systems prior to site visit (allowing for adequate time for site contact to obtain).

Outcome: Improved data collection and customer satisfaction.

ENBRIDGE RESPONSE: Enbridge is unaware of any customer site visits/project reviews where verifiers indicated any concerns with the conduct of Enbridge utility staff, however as communicated to the EC early in the process, Enbridge shared concerns regarding observations of the verifiers at a number of the site visits. These included poor/untimely communication regarding site visit scheduling, concerns about questions asked of customers regarding unrelated or irrelevant information about the project indicating a poor understanding of the project or technology, site visit reports that included measurements or findings that were in fact not completed, and requests for data that were perceived to compromise customer privacy.

Enbridge concurs that a verification code of conduct for verification and utility staff should be established. Enbridge also suggests that protocols ensure there is a project review with utility staff undertaken prior to the site visit to ensure a clear understanding of the project.

Table 5-8 Documentation and Support recommendations[13]

#	Documentation and Support Recommendation	Applies to		
		Union	Enbridge	Evaluation
DS1	Take steps to improve documentation: • Include explicit sources for all inputs and assumptions in the project documentation. • Store background studies and information sources with the project files and make them available to evaluators. • Provide evaluators full access to customer data. • Provide pre- and post-installation photos, where available. • Document and provide internal M&V documents where available. • Institute a checklist as part of project closeout to ensure all relevant project documentation is assembled as ready for verification	✓	✓	
DS2	Ensure that incremental costs are supported by invoices or other documentation	✓	✓	
DS3	Increase the amount of documentation and source material for projects that have greater energy savings.	✓	✓	

[13] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

DS4 A	Digitize and file project documentation for all projects as they are completed and paid during project closeout.	✓	✓	
DS4 B	Until the utilities can implement an effective digital document storage process, the evaluation should allow more time for the utilities to assemble and deliver the documentation.			✓
DS5	Consider providing more training or adding quality control steps to ensure the summary workbook front page is completed and stored in a consistent manner.	✓		
DS6	Use a consistent summary workbook.		✓	

DS1. Finding: Project documentation for some projects lacked sufficient details to allow evaluators to reproduce the calculations made by program staff or third-party vendors. Specific issues included:

- Project data or details missing
- Insufficient measure-level details to fully describe what was installed
- Descriptions that were difficult to understand
- Use of black box tools
- Hardcoded information in calculation spreadsheets
- Energy intensity changes presented without providing the data to justify it
- Undocumented assumptions
- Sources referenced but not included or available, such as feasibility studies and historical analysis of energy use that was left out of the project documentation
- Scanned documents that were unreadable
- Input adjustments that approximate other effects, but are not explained
- Insufficient access to customer data (by customers) for confidentiality reasons.

- Modelling files that could not be opened
- Adjustments to savings estimates for safety or influence that were not clearly marked, sourced, or carried out in a consistent fashion
- Etools files not provided for many industrial boiler & boiler add-on projects

Recommendation: Several steps could be taken to improve data quality:

- Include explicit sources for all inputs and assumptions in the project documentation.
- Store background studies and information sources with the project files and make them available to evaluators.
- Provide evaluators full access to customer data.
- Provide pre- and post-installation photos, where available.
- Document and provide internal M&V documents where available.
- Institute a checklist as part of project closeout to ensure all relevant project documentation is assembled as ready for verification

Outcome: Properly explaining and sourcing the savings calculation method and assumptions allows the evaluating engineer to more easily identify what needs to be verified. It also makes it easier to determine whether the methods and assumptions are reasonable and use ex ante assumptions rather than seek documented values elsewhere.

ENBRIDGE RESPONSE: Enbridge continually strives to improve the comprehensiveness of custom project documentation and generally works to ensure full and detailed inputs and supporting evidence is clearly outlined for each project. Nonetheless, Enbridge will review these recommendations to improve data quality moving forward. Given the timing of the receipt of the 2015 Annual Verification

Recommendations, in Q3 of 2017, incorporation of any such recommendations will be made in the 2018 program year.

DS2. Finding: Invoices were not always included with documentation, and we saw a handful (<5) of cases where utility program staff were overclaiming incremental costs. This did not appear to be systemic, but higher incremental costs enable payment of a larger incentive.

Recommendation: Ensure that incremental costs are supported by invoices or other documentation, especially for add-on and optimization measures where the total cost and incremental cost are likely to be the same. Equipment replacement measures may require an additional standard efficiency quote to produce incremental cost.

Outcome: Incremental cost is an important component of simple payback, which is often used to judge the economic benefit of energy efficiency projects. It is also an input to some benefit-cost tests.

ENBRIDGE RESPONSE: Enbridge endeavours to ensure that claimed incremental costs are supported by invoices or other documentation. In some cases, project costs are bundled within invoices for larger work being completed in tandem at a customer site. Enbridge will continue to work to minimize any instances where incremental costs are not clearly documented.

DS3. Finding: Larger projects appeared to fall under the same documentation standards as smaller projects.

Recommendation: Increase the amount of documentation and source material for projects that have greater energy savings.

Outcome: Projects that are better documented tend to have more accurate savings estimates and receive fewer evaluation adjustments than those that are less documented. Large projects have a greater effect on overall savings adjustment factors. Therefore, large projects with better documentation are more likely to result in adjustment factors closer to 100%.

ENBRIDGE RESPONSE: For consistency, Enbridge strives to ensure project documentation captures all the relevant information to support accurate savings calculation estimates regardless of the size of project.

DS4. Finding: Enbridge did not maintain complete digital project files prior to the evaluation request. Union appeared to have digital documentation that was not completely assembled prior to evaluation.

Recommendation A: Digitize and file project documentation for all projects as they are completed and paid during project closeout. PDF and Excel files associated with a project should be stored in a way that allows them to be easily found and associated with a specific project and/or customer. The best practice is to include a document repository as part of the program tracking system with a separate folder for each project.

ENBRIDGE RESPONSE: Enbridge DSM is currently undergoing a DSM IT system upgrade that will include improvements to the organization and facilitation of digitized project files. This system upgrade is expected to be rolled out in 2018.

Recommendation B: Until the utilities can implement an effective digital document storage process, the evaluation should allow more time for the utilities to assemble and deliver the documentation.

Outcome: In our experience, DSM programs that store complete and well-organized digital records experience less evaluation risk. In other words, their gross savings adjustments are closer to 100%. This happens for three reasons:

- Digitization facilitates internal review of project documentation, providing additional opportunities to identify missing information and errors
- Assembly during project closeout improves the comprehensiveness of the documentation because less time has elapsed than if it was assembled for evaluation, so less information is lost or forgotten

Easy retrieval makes it more likely that the complete file is sent to the evaluation team, reducing the information gap between implementation and evaluation.

<u>ENBRIDGE RESPONSE</u>: This recommendation was not directed to Enbridge

DS5. Finding: Union custom projects utilized a project application summary workbook that summarizes the key project inputs, calculations, and most details. In general, this is a good approach that facilitates internal review and evaluation. One challenge was that different projects used the workbook in different ways:

- The notes section was sometimes used to identify and highlight specific unique approaches and features in projects, but not always.
- Calculations internal to the summary page were consistent for most projects, but not all (additional factors were sometimes added).
- Sub-methods critical to the calculation were contained in hidden sheets.
- Safety and influence adjustments were inserted in different locations and not always explained.

Recommendation: Consider providing more training or adding quality control steps to ensure the summary workbook front page is completed and stored in a consistent manner. Identify a common approach for common measures and, if necessary, document deviations and the reasons for the deviations in a clearly labelled field on the summary sheet.

Outcome: A consistent summary workbook aids both internal and external quality assurance, quality control, and measurement and verification.

ENBRIDGE RESPONSE: Not Applicable

DS6. Finding: The Enbridge Etools is used as both a calculation tool and as a communication tool with customers. While it appears to serve the needs of the program, this form of communication is difficult for the evaluation efforts.

- Etools does not easily allow for assumptions to be sourced within the record.
- Some Etools selections may be site-specific and some may be defaults; the calculator does not distinguish.
- Energy savings that are calculated outside of Etools are hard-entered in Etools but not always sourced.

Recommendation: Use a consistent summary workbook.

Outcome: A consistent summary workbook aids both internal and external quality assurance, quality control, and measurement and verification.

ENBRIDGE RESPONSE: Enbridge is committed to continue in its efforts to improve upon the comprehensiveness and clarity of all relevant project information, data and underlying input assumptions. Given the timing of the receipt of the 2015 Annual

Verification Recommendations, in Q3 of 2017, considerations to improve on a project summary workbook will be reviewed for the 2018 program year.

Table 5 9: Data management recommendations[14]

#	Data Management Recommendation	Applies to		
		Union	Enbridge	Evaluation
DM1 A	Track contacts associated with projects in the program tracking database.	✓	✓	
DM1 B	Strongly consider investing in relational program tracking databases.	✓	✓	
DM1 C	Include structure for improved data integrity in the evaluator request for contact information for the 2016 and 2017 savings verification and evaluation.			✓
DM2 A	Consider offering bonus incentives early in the year to combat the "hockey stick" phenomenon where a large percent of projects get closed in the fourth quarter of the year (which results in rushed QC for data).	✓	✓	
DM3	Track and provide to evaluators dates for key milestones in the project.	✓	✓	
DM4	Maintain a customer identifier in the database to clearly identify related sites.	✓	✓	
DM5	Include EUL (also remaining useful life for dual baselines), NTG, and each of the key savings types (i.e., annual and cumulative, gross and net) in the program tracking extracts provided to evaluators.	✓	✓	

[14] EB-2015-0245, Ontario Gas DSM Evaluation Contractor, 2015 Natural Gas Demand Side Management Annual Verification, October 12, 2017

DM1. Finding: Neither Union nor Enbridge currently track participating customer or participating vendor contact information in their program tracking database. Providing the information to the evaluation put significant burden on utility staff. When contact information was provided, there were significant data integrity issues including contacts listed in the wrong places, partial addresses, and incorrect or missing phone numbers and email addresses.

Recommendation A: Track contacts associated with projects in the program tracking database. At a minimum, the program tracking database should include:

- Project site address
- Customer mailing address
- Primary customer contact name
- Primary customer contact phone
- Primary customer contact email
- Primary customer contact mailing address
- Addresses are best tracked as multiple fields including:
    - Street address line 1
    - Street address line 2
    - City
    - Province
    - Postal code

Phone number fields should include data validation to enforce a consistent format and avoid missing or extra digit errors. Phone extensions should be tracked in a field separate from the ten-digit phone number and be restricted to numeric data only.
The best practice is to maintain contacts in a table separate from specific project or customer data. This allows for a single contact to be connected to multiple accounts and/or projects as necessary without creating duplication. This structure also makes it easier to associate multiple contacts with a single project.

Vendor contact information should also be tracked in the database, in the same table as the participating customer contact information. With a relational database, the contact ID from the table can be added to a project record in the role consistent with the contact's participation (such as vendor, decision maker, or technical expert) with a separate table that allows a single vendor contact to be associated with multiple projects.

Outcome A: Reduced burden on utility staff to seek contact information for projects, whether for internal or evaluation use. Reduced evaluation costs and improved sample design expectations.

ENBRIDGE RESPONSE: Enbridge is currently undergoing a tracking & reporting system upgrade that will enable the capture of participant and vendor information in a single database. This upgrade is expected to be rolled out in 2018.

Recommendation B: The utilities should strongly consider investing in relational program tracking databases. Relational program tracking databases and customer relationship management ("CRM") systems allow for multiple contacts to be associated with a single account and/or project. This allows programs to easily clarify aspects of projects during implementation and to provide accurate, timely, and usable contact information to evaluators and verifiers. The incremental cost of implementation is low if it is part of the initial database design, populated as projects are started, and updated once they are complete.

Outcome B: Reduced burden on utility staff and reduced evaluation costs. A relational database would streamline aggregation of program data for scorecards and make providing data simpler for annual savings evaluation and verification.

ENBRIDGE RESPONSE: Enbridge DSM is currently undergoing an IT system upgrade that will include improved tracking & reporting and CRM components. This system upgrade is expected to be rolled out in 2018.

Recommendation C: For 2016 (and perhaps 2017), we do not anticipate that contact information will have been entered into the program tracking databases. When the evaluation requests contact information for the 2016 and 2017 savings verification and evaluation, the contact request spreadsheet will be updated to provide additional fields to enforce data integrity (e.g., specific fields for a parsed address and company name for the technical and decision-making contacts).

Outcome C: Reduced evaluation costs due to less data cleaning and research to fill missing information. Improved data collection with less returned advance letters and more accurate connection between projects and contacts.

ENBRIDGE RESPONSE: This recommendation was not directed to Enbridge

DM2. Finding: Both utilities have indicated that inputting and/or extracting data necessary for annual reporting and evaluation requires significant effort.

Recommendation A: Consider offering bonus incentives early in the year to combat the “hockey stick” phenomenon where a large percent of projects get closed in the fourth quarter of the year.

Outcome: Reduced burden on program staff, more consistency in meeting annual filing deadlines.

<u>ENBRIDGE RESPONSE</u>: In the case of the 2015 Verification, the data requests from the EC were delivered in Q4 at year end, not in Q2 or Q3 as contemplated in the Board's new governance structure. This is the busiest time of the year. However, Enbridge does not agree that offering a bonus incentive early in the year to combat the "hockey stick" phenomenon would address the EC's finding. Firstly, in many cases, particularly in industrial setting, customers utilize primarily two time periods to execute a major change to their process or the facility: summer shutdown, for those customers that incorporate this mid-year break and more often Christmas shutdown.

The EC's observation regarding the utilities' effort was in large part as a result of the change in process, new data categorization requirements and the increased volume of project files requested for the CPSV, free-ridership and spillover studies encompassing substantially more projects compared to previous years.

Recommendation B: See recommendation DM1B. The utilities should consider investing in a new database.

Outcome: Reduced burden on utility staff and reduced evaluation costs.

<u>ENBRIDGE RESPONSE</u>: See Enbridge's response to DM1B.

DM3. Finding: The extracts from the utility program tracking database do not include dates for key project milestones. Enbridge's data did not include any dates and Union's included only the "installation date."

Recommendation: Track and provide to evaluators dates for key milestones in the project.

Dates for project start, installation, and those that define the program year provide useful context for interviewers that is not always easy to find in project documentation

Outcome: Improved data collection through more informed interviewers and reduced evaluation costs through less need to search for dates in documentation.

ENBRIDGE RESPONSE: Though all "key project milestones" may not have been presented in Enbridge's summary tracking database, Enbridge does include the measure(s) "Turn on Date", which denotes when the measure(s) has been installed and fully commissioned in the tracking summary. The incentive payment process commences only after this date has been entered into the tracking database and the custom project file submitted for claim. This date is also utilized for LRAM purposes. In addition ESCs are expected to track other relevant key milestones in the project file including project initiation and meeting dates with customers.

Enbridge DSM is currently undergoing an IT system upgrade that will included improved tracking & reporting and CRM components that will facilitate the improved capture of milestone dates. This upgrade is expected to be rolled out in 2018.

DM4. Finding: Customers with multiple sites are not tracked in the program tracking database. A few property management groups had many sites selected in the sample, but it was not clear from project tracking or the provided contact information that the sites were related. Property management firms were the most significant but not the only customer type where this was true.

Recommendation: Maintain a customer identifier in the database to clearly identify related sites. This is easiest to deploy in a relational database see recommendation DM1B.

Outcome: Reduced evaluation costs and reduced customer burden. In some cases, a failure to identify related sites can result in multiple calls to the same customer, which a customer identifier would avoid. In addition, tracking related sites could improve program implementation by increasing awareness of connected opportunities.

ENBRIDGE RESPONSE: Enbridge's custom projects are designated with a unique project ID. Although a customer identifier to identify related sites is not utilized for custom projects, projects can be linked on the basis of billing information, site address, or at the customer name assignment for multiple addresses. There are some exceptions however are schools boards and property managers with many sites.

DM5. Finding: EUL and cumulative gross savings were not provided in the standard program tracking database extracts. The evaluation team backed out the missing information from the fields provided.

Recommendation: Include EUL (also remaining useful life for dual baselines), NTG, and each of the key savings types (i.e., annual and cumulative, gross and net) in the program tracking database.

Outcome: Improved data integrity results in less evaluation risk and more accurate savings totals. Providing each of the key savings types and their components allows evaluation to confirm that the savings provided are internally consistent.

ENBRIDGE RESPONSE: Enbridge tracks the EUL for all custom projects and includes the RUL where it is determined to be applicable; in addition Enbridge includes the remaining categories listed above in its tracking summary.

Ontario Energy Board

P.O. Box 2319
2300 Yonge Street
27th Floor
Toronto ON M4P 1E4
Telephone: 416-481-1967
Facsimile: 416-440-7656
Toll free: 1-888-632-6273

Commission de l'énergie de l'Ontario

C.P. 2319
2300, rue Yonge
27e étage
Toronto ON M4P 1E4
Téléphone: 416-481-1967
Télécopieur: 416-440-7656
Numéro sans frais: 1-888-632-6273

Ontario

BY EMAIL AND WEB POSTING

August 21, 2015

To: **All Natural Gas Distributors**
All Participants in the Consultation Process EB-2014-0134
Other Stakeholders

Re: **2015-2020 Demand Side Management Evaluation Process of Program Results**
EB-2015-0245

This letter establishes the OEB's process to evaluate the results of Natural Gas Demand Side Management (DSM) programs from 2015 to 2020.

Background

As outlined in Section 7 of the OEB's Report on DSM issued December 22, 2014, the OEB will be taking a central role in the evaluation process of DSM program results. DSM programs will be evaluated on an annual basis, with results issued by the OEB to be used by the gas utilities when they file applications for recovery of amounts related to DSM activities.

DSM Evaluation Governance

The OEB will rely on the DSM evaluation governance structure outlined below. The evaluation governance structure describes the general role of the main parties involved in the evaluation process. The evaluation governance structure is expected to be fully implemented following the OEB's selection of an Evaluation Contractor.

OEB's DSM Evaluation Governance Structure	
Party	**Role**
OEB	The OEB is responsible for coordinating and overseeing the evaluation and audit process, including selecting a third party Evaluation Contractor and publishing the final evaluation results on an annual basis.
Evaluation Contractor (EC)	The Evaluation Contractor will carry out the evaluation and audit processes of all DSM programs.
Natural Gas Utilities	The natural gas utilities are responsible for developing an initial evaluation plan that will inform the evaluation of programs, filing an annual draft evaluation report and providing program data and coordination support to the Evaluation Contractor and OEB staff, as requested.
Evaluation Advisory Committee (EAC)	An Evaluation Advisory Committee (EAC) will be formed to provide input and advice to the OEB on the evaluation and audit of DSM results. The EAC will consist of representatives from non-utility stakeholders, independent experts, staff from the Independent Electricity System Operator (IESO), and observers from the Environmental Commissioner of Ontario and the Ministry of Energy, all working with OEB staff.

Evaluation Approach

The OEB will retain a third party Evaluation Contractor to undertake DSM program evaluations and annual audits of program results.

The Evaluation Contractor will draft an Evaluation, Measurement & Verification (EM&V) Plan for the natural gas utilities' DSM programs for approval by the OEB. The EAC will provide advice and input on the development of the plan as required. The EM&V Plan will, at a minimum, address the following:

- Annual Evaluation and Audit of DSM results
- Annual update of input assumptions
- Multi-year DSM program impact assessments and evaluation studies

The OEB-approved EM&V plan is expected to span a period of three-years to coincide with the mid-term review of both the 2015 to 2020 Natural Gas DSM Framework and Electricity CDM Framework.

Annual Evaluation & Audit Process

Consistent with current evaluation practices, the Evaluation Contractor will be responsible for auditing each gas utility's annual DSM results based on the three-year OEB-approved EM&V plan. The detailed annual evaluation and audit process will be developed as part of the EM&V plan.

Updating Input Assumptions

The Evaluation Contractor will review and propose updates to the OEB related to data within the Technical Reference Manual (TRM) on an annual basis. This review of the TRM will include proposed updates to input assumptions to reflect the findings of the annual DSM evaluation and audit. This may require additional research in order to add any new technologies to the TRM and improve the current list of assumptions.

Best efforts will be made to align the natural gas DSM input assumptions list with the electricity CDM input assumption list, where appropriate. The OEB is of the view that having alignment on resource savings amounts related to both natural gas and electricity energy efficiency technologies will help enable a greater level of integrated and collaborative program design and delivery.

Multi-Year DSM Program Impact Assessments and Evaluations

The OEB will engage the Evaluation Contractor to conduct multi-year impact assessments and targeted evaluations of selected natural gas DSM programs on a periodic basis throughout the 2015 to 2020 DSM period.

Within the Evaluation Contractor's multi-year impact assessments, the Evaluation Contractor will be responsible for undertaking various studies which may include estimating natural gas savings, undertaking net-to-gross studies, investigating free ridership rates and spillover effects, examining the level of persisting natural gas savings from various programs and conducting other evaluation studies as required.

Transition Plan

The OEB recognizes that there is a current evaluation process underway, led by the natural gas utilities with support from three committees: the Technical Evaluation Committee (TEC), and two Audit Committees (one for each utility). The committees are comprised of natural gas utility staff, industry stakeholders and independent experts.

The current responsibilities of the TEC include the development of the Technical Reference Manual (TRM), the completion of a Commercial and Industrial Custom Project Net-to-Gross Study, a joint utility Boiler Baseline Study, and the initiation of a Persistence Study. This is important work that should continue at this time. The evaluation and audit of all natural gas DSM program results under the new 2015–2020 DSM Framework will follow the new process outlined in this letter. Once an Evaluation Contractor is retained by the OEB, OEB staff will work with the TEC on an appropriate plan to transition to the new framework on a go-forward basis. With the formation of an Evaluation Advisory Committee (EAC), as described below, an Audit Committee will no longer be required.

Formation of the Evaluation Advisory Committee

The Evaluation Advisory Committee (EAC) will provide input and advice as required throughout the DSM evaluation process. The EAC will be comprised of:

- Experts representing non-utility stakeholders, with demonstrated experience and expertise in the evaluation of DSM technologies and programs, natural gas energy efficiency technologies, multi-year impact assessments, net-to-gross studies, free ridership analysis and natural gas energy efficiency persistence analysis
- Expert(s) retained by the OEB
- Representatives from the IESO
- Representatives from each natural gas utility
- Representatives from the Ministry of Energy (MOE) and the Environmental Commissioner of Ontario (ECO), who will participate as observers

The OEB has recently selected a group of experts representing non-utility stakeholders to provide input and advice as part of the DSM Technical Working Group formed for the natural gas conservation potential study. As the technical expertise and experience required for both the DSM Technical Working Group and EAC are similar, the OEB has appointed the same individuals to represent non-utility stakeholders on the EAC as follows:

- Chris Neme, Energy Futures Group
- Jay Shepherd, Jay Shepherd Professional Corporation
- Marion Fraser, Fraser & Company

Due to a potential conflict, Ian Jarvis, who is a member of the DSM Technical Working Group, has not been included as a member of the EAC.

In reviewing nominations from non-utility stakeholders as part of the formation of the DSM Technical Working Group, the OEB considered the diversity of their expertise, their participation in similar OEB proceedings and working groups and their experience with the Ontario natural gas sector, as well as their ability to represent stakeholders. The selected candidates are expected to provide input and advice based on their experience and technical expertise and not to advocate position of parties they have represented before the OEB in various proceedings.

The OEB will determine the appointment of additional experts following the selection of an Evaluation Contractor.

Cost Awards

Cost awards will be available under Section 30 of the *Ontario Energy Board Act, 1998* to eligible persons in relation to their participation in the Evaluation Advisory Committee or other consultations during the course of the DSM evaluation process. Details will be provided at the appropriate time. Costs awarded will be recovered from all rate-regulated natural gas distributors based on their respective distribution revenues.

If you have any questions regarding this consultation process, please contact Josh Wasylyk at Josh.Wasylyk@OntarioEnergyBoard.ca or at 416-440-7723.

The OEB's toll free number is 1-888-632-6273.

Yours truly,

Original Signed By

Kirsten Walli
Board Secretary

Ontario Energy Board

P.O. Box 2319
2300 Yonge Street
27th Floor
Toronto ON M4P 1E4
Telephone: 416-481-1967
Facsimile: 416-440-7656
Toll free: 1-888-632-6273

Commission de l'énergie de l'Ontario

C.P. 2319
2300, rue Yonge
27e étage
Toronto ON M4P 1E4
Téléphone: 416-481-1967
Télécopieur: 416-440-7656
Numéro sans frais: 1-888-632-6273

Ontario

MEMORANDUM

BY EMAIL

March 4, 2016

To: **Enbridge Distribution Inc.**
Union Gas Limited
Technical Evaluation Committee (TEC)
Evaluation Advisory Committee (EAC)

Re: **Transition of Technical Evaluation Committee Activities to the OEB**
EB-2015-0245

The purpose of this memorandum is to map out the transition of the current demand-side management (DSM) evaluation activities from the TEC to the OEB.

Background

As described in the OEB's August 21, 2015 letter regarding the 2015-2020 DSM evaluation process, the TEC's evaluation activities will be transitioned to the OEB under the new DSM evaluation governance structure. In that letter, the OEB directed the gas utilities and the TEC to continue working on the evaluation projects that they had initiated until the transition takes place. The current status of each of the Projects is as follows:

1) *Technical Reference Manual (TRM) Development*: The TEC selected two vendors to complete the TRM: Energy & Resource Solutions (ERS) and MindTouch, Inc. ERS was procured to complete the first phase of the project (development of the TRM with updated measures and input assumptions). This project is mostly completed. MindTouch was procured to complete the second phase (online platform of TRM) of the project.

2) *Net-to-Gross Study:* The Custom Commercial and Industrial Net-to-Gross Study will update the current net-to-gross rates used to estimate the impact of custom projects. The utilities, with the endorsement of the TEC, managed the procurement process and contracted DNV KEMA (now DNV GL) in May 2015 to undertake the study. DNV GL is in the process of developing a detailed Work Plan.

3) *Boiler Baseline Study:* As part of separate OEB Decisions issued in 2015 (February 26, 2015[1] and June 4, 2015[2]), Union and Enbridge were directed to complete a Boiler Baseline Study in 2015, with the findings incorporated in the evaluation of 2014 results. The TEC developed a study scope of work and issued a Request for Proposals (RFP) for this study in October 2015. In response to the RFP, consultants have submitted proposals to the TEC. The TEC has not proceeded with the evaluation of the proposals as it is awaiting further instructions from the OEB.

4) *Persistence Study*: The TEC conducted initial research into the scope of work for a persistence study in 2015. This study has not been initiated.

Transitioning to the OEB

The transition plan for each study is outlined below.

1) *Technical Reference Manual (TRM) Development*

Development of the TRM with updated measures and input assumptions is mostly completed and the TEC will continue to finalize the TRM with ERS. The management of the online portion of the TRM has been transitioned to OEB Staff, who will post the final TRM online when it is available. The utilities will continue to manage any remaining contractual obligations and payments related to the TRM.

2) *Net-to-Gross Study*

DNV GL plans to present the draft work plan for the net-to-gross study at the next TEC meeting, currently scheduled for March 10, 2016. Following input from the TEC, this

1 EB-2014-0277 – Enbridge Gas Distribution Inc.
2 EB-2014-0273 – Union Gas Limited

study will be transitioned to OEB. The utilities will continue to manage contractual obligations and payments associated with this project. OEB Staff will assume oversight of the study and will confirm the completion of major milestones for the utilities to process payments of consultant's invoices.

Though OEB Staff will have oversight going forward for the TRM and Net-to-Gross Study as noted above, the gas utilities will incur the costs to complete these studies and therefore can seek recovery of these costs as part of the DSM program.

3) *Boiler Baseline Study*

This study was the result of OEB decisions for both Enbridge and Union Gas and therefore the utilities are expected to complete it. Once the proposals have been evaluated and the consultant selected for the Boiler Baseline Study, in order to transition to the new framework, input on the study will be provided to the utilities by the EAC and OEB Staff instead of the TEC.

4) *Persistence Study*

OEB Staff will be responsible for the procurement process and management of the Persistence Study, including management of project deliverables and contractual obligations through to completion of the study, with input from the EAC.

If you have any questions regarding the transition process, please contact Takis Plagiannakos at takis.plagiannakos@ontarioenergyboard.ca or 416-440-7680.

Yours truly,

Original Signed by

Lynne Anderson
Vice President, Applications

PRICING INFORMATION REMOVED

MEASUREMENT OF NTG FACTORS FOR ONTARIO'S NATURAL GAS CUSTOM COMMERCIAL AND INDUSTRIAL DSM

Scope of Work

for Ontario Natural Gas Technical Evaluation Committee (TEC)

Date: 3/2/2016

Table of Contents

Tables, Figures, Equations

OVERVIEW

This document presents the scope of work for the measurement of net-to-gross (NTG) Factors for Ontario's Natural Gas Custom Commercial and Industrial demand side management (DSM) programs for the Ontario Natural Gas Technical Evaluation Committee (TEC).

The two largest gas utilities in Ontario, Union Gas Limited (Union) and Enbridge Gas Distribution (Enbridge), (together, the "utilities") have offered DSM incentives to businesses for implementing energy efficiency improvements for twenty years. The Union custom incentives are provided as part of the Union commercial program and as part of the direct access program for large industrial customers. The Enbridge custom incentives are provided as part of the Enbridge commercial and industrial programs as well as its Run-it-Right retro-commissioning program.

This evaluation will assess the NTG factors for custom measures in the Union large industrial and commercial programs and the Enbridge commercial, industrial, and Run-it-Right programs. This work plan is a living document that will be updated as new data is incorporated and additional decisions are made.

Evaluation Objectives

The overall goal of this evaluation is to develop transparent free ridership and spillover factors for custom commercial and industrial programs, to be used for future programs.

Evaluation Approach

The methodology selected for this evaluation will rely on end-user self-report surveys and interviews to estimate program NTG. The end user self-reports will be supplemented by project-specific interviews with vendors and vendors to capture indirect effects of the program on end-user decision making. Surveys and interviews will be collected from the most recent program years in order to create NTG factors that will be most meaningful for future years.

For Union's large industrial program and the largest commercial projects and the largest Enbridge industrial and commercial projects, we will estimate NTG using participating end user self-reports and project-specific interviews with vendors.

Key Concepts

This section defines several key concepts that will be used throughout this work plan, using the definitions from the Ontario DSM Guidelines for spillover and free rider.

- *Spillover* "refers to effects of customers that adopt energy efficiency measures because they are influenced by a utility's program-related information and marketing efforts, but do not actually participate in the program." [1] We consider both inside and outside spillover through this project.
    - *Inside spillover* "refers to non-incented measures that were installed within the same project or facility."[2]
    - *Outside spillover* "refers to measures for which the customer did not receive an incentive adopted in an outside location or unrelated project for a participating customer." [3]

[1] Ontario Energy Board *Demand Side Management Guidelines for Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

[2] Ontario Natural Gas Technical Evaluation Committee (TEC), *Request for Proposal: Measurement of Net-to-Gross (NTG) Factors for Ontario's Natural Gas Custom Commercial and Industrial Demand Side Management (DSM) Programs,* RFP-002-2013 (2), December 2013, Section 2.

- A *free rider* is "a program participant who would have installed a measure on his or her own initiative even without the program." [4]
- *Gross savings* are "the changes in energy consumption and/or demand that result directly from program-related actions taken by participants in an efficiency program, regardless of why they participated."[5]
- *Net savings* are "the changes in energy consumption or demand that are attributable to an energy efficiency program. The primary, but not exclusive, considerations that account for the difference between net and gross savings are free riders (i.e., those who would have implemented the same or similar efficiency projects, to one degree or another, without the program now or in the near future) and participant and non-participant spillover (i.e., savings that result from actions taken as a result of a program's influence but which are not directly subsidized or required by the program). Net savings may also include consideration of market effects (changes in the structure of a market)."[6]
- The *net-to-gross* (NTG) ratio is an adjustment factor that reduces savings due to free ridership and increases savings to account for spillover. The NTG ratio "is the portion (it can be less than or greater than 1.0) of gross savings (those that occur irrespective of whether they are caused by the program or not) that are attributed to the program being evaluated."[7]

---

[3] Ontario Natural Gas Technical Evaluation Committee (TEC), *Request for Proposal: Measurement of Net-to-Gross (NTG) Factors for Ontario's Natural Gas Custom Commercial and Industrial Demand Side Management (DSM) Programs,* RFP-002-2013 (2), December 2013, Section 2.

[4] Ontario Energy Board *Demand Side Management Guidelines for Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

[5] SEE Action, Energy Efficiency Program Impact Evaluation Guide: Evaluation, Measurement, and Verification Working Group, DOE/EE-0829, December 2012. https://www4.eere.energy.gov/seeaction/sites/default/files/pdfs/emv_ee_program_impact_guide_1.pdf, page xiv

[6] Ibid

[7] Ibid, page 5-1

SUMMARY OF TASKS

The DNV GL team has broken the project into six discrete tasks which are presented, along with their status, in Table 1. These tasks are discussed in greater detail in the next sections of the report.

Table 1: Key Project Tasks and Completion Status

Key Tasks
☒ **Task 1: Project Kickoff**
☒ Convene a project kickoff meeting
☒ Reach a consensus on methodology
☐ **Task 2: Sample Design**
☐ Explore the tracking data
☐ Define the unit of analysis
☐ Stratify the data
☐ Design the sample
☐ Prepare the sample frame
☐ **Task 3: Data Collection**
☒ Interview program managers and staff
☐ Interview program Energy Solution Consultants (ESC)
☐ Survey program participants
☐ Interview large or complex program participants
☐ Interview program trade allies
☐ Conduct follow-up interviews with program participants
☐ **Task 4: Data Analysis**
☐ Analyze survey and interview data
☐ Calculate estimates
☐ **Task 5: Reporting**
☐ Produce an evaluation report identifying free ridership and spillover factors for custom commercial and industrial programs
☐ **Task 6: Project Management**
☐ Complete evaluation on time, on budget and within scope
☐ Keep client informed on progress

We have completed the project kickoff meeting, program manager and staff interviews, and initial sample design as part of the planning phase, which have informed the specific plan outlined in this document. Once DNV GL receives the complete tracking dataset we will create the full sample design which will be used to select projects for computer aided telephone interview (CATI) surveys and expert in-depth interviews (IDI). Next, we will request the contact information and necessary documentation to proceed to the participant data collection phase. This will also include interviews with Energy Solution Consultants (ESCs) and vendors who have completed projects through the program. We will calculate the free ridership, spillover, and NTG estimates for each program and domains within programs where there is sufficient sample to provide estimates while protecting respondent confidentiality. These estimates will be provided in the final evaluation report.

Task 1: Project Kickoff

Meeting and Follow-up Memorandum Overview

The kickoff meeting on March 17, 2014 was an in-person meeting between the TEC and Evaluation leadership. Most of the time was spent on high level evaluation concepts and in understanding different perspectives within the TEC. This meeting was followed by a series of memos with project decisions on April 1, 2014 (kickoff summary), June 12, 2014 (kickoff parking lot items), and July 2, 2014 (methodology explanation). The kickoff was held prior to contract negotiation, which was completed in mid-2015.

Table 2: Task 1 Tasks and Completion Status

Task 1: Project Kickoff Subtasks
☒ **Task 1.1: Convene a project kickoff meeting**
☒ **Task 1.2: Reach a consensus on methodology**

Resolved Items

- **Consensus around primary project objective.** The primary objective of this project is a transparent, reputable study that produces strong, credible, and defensible NTG ratios to be used on a go-forward basis. (4/1/2014)

- **Concern about scope creep.** The potential for scope creep, particularly analysis and reporting of information collected, but not part of the NTG estimation, is a concern of several members of the TEC and evaluation team. (4/1/2014)

- **Consensus for TEC review of data collection instruments.** DNV GL will submit survey instruments, along with probes, question rationale and scoring to the TEC. (6/12/2014)

- **Consensus for qualitative reporting of participant decision making reasons.** DNV GL will include qualitative discussion of participant-reported reasons for results describing NTG and spillover analysis results. (6/12/2014)

- **Stratification determined by DNV GL.** Stratification of survey participants will be representative, as determined by DNV GL's expert judgment. (6/12/2014)

- **First Year Net Savings (Y1NS) method recommended.** DNV GL recommends the use of the Y1NS method for the current NTG study. The LCNS method requires engineering calculations that would add additional scope to the standalone NTG study. (7/2/2014).[8]

[8] July 2, 2014 DNV GL Memo to TEC: *Attribution Method Comparison (Y1NS vs LCNS)*.

Task 2: Sample Design

The objective of the sample design is to select customers for surveys and IDIs to estimate the free ridership and spillover for the custom C&I projects and to create an optimized plan for data collection and expansion. Prior to completing the sample design, we determined that we are likely to attempt a census of participants due to the ratio of targeted completes to accounts in the data provided. Even though we intend to attempt a census, we completed most of the steps required in a sample design to have a basis for post-stratifying the completed surveys and IDIs for expansion to the population.

Through the sample design process, we define:

- The unit of analysis
- The number of surveys targeted for each program
- The number of IDIs targeted for each program
- The stratification that will be used for expansion

This section presents the stratification plan using the initial datasets for 2013 and 2014 custom C&I projects provided by Union and Enbridge. We anticipate receiving updated data, including 2015 projects and the Run-it-Right project data in early March. Table 3 presents the sample design tasks and their completion status.

Table 3: Task 2 Subtasks and Completion Status

Task 2: Sample Design Subtasks
☐ **Task 2.1: Explore the Tracking Data**
☒ Initial data exploration , Union and Enbridge
☐ Exploration of the full datasets, Union and Enbridge
☐ **Task 2.2: Define the Unit of Analysis**
☒ Initial unit of analysis definition
☐ Final unit of analysis definition using full datasets
☐ **Task 2.3: Stratify the Data**
☒ Initial stratification
☐ Final stratification using full datasets
☐ **Task 2.4: Design the Sample**
☒ Initial sample design
☐ Full sample design and precision optimization
☐ **Task 2.5: Prepare the Sample Frame**
☐ Sample contact information and documentation request
☐ Prepare the sample frame

Task 2.1: Explore the Tracking Data

We first explored the tracking data provided by Union and Enbridge to determine data availability, the number and types of measures installed, as well as the size and quantity of projects. We explored the Union and Enbridge datasets separately.

Enbridge Participant Data

The participant data files provided by Enbridge included custom C&I energy efficiency projects completed during the 2013 and 2014 program years (Table 4). DNV GL has not yet received the 2015 tracking data, nor data for the Run-it-Right program. These data will be included in the final version of this plan. In the 2013 and 2014 Enbridge tracking data, there are 1,603 records and 1,189 unique accounts. The records in the tracking data are per installed measure, so there may be multiple rows per project if more than one measure is installed, and there may also be more than one project per account. Across the three program years, we expect to have approximately 2,400 records for approximately 1,600 unique accounts.

Table 4: Enbridge Participation Metrics by Year

Program Year	Accounts	Gas Savings (m^3)
2013	681	53,030,333
2014	576	46,195,015
2015 projection*	576	46,195,015

*These data are duplicates of 2014 for reference purposes only.

The tracking data includes measure level savings specific to a measure, site and date. As part of defining the unit of analysis, we used the tracking data variables *load type name*, *end use*, and *technology* to categorize measures into measure types (Table 5). The tracking data we received had some conflicts among identifiers that made it difficult to create consistent measure types across the sample frame. For our initial sample design we divided the measures into two categories: equipment and operations & maintenance (O&M), but we plan to revisit this decision once we receive the complete dataset and we know the overall distribution of measures.

Table 5: Approximate Enbridge Participation Metrics by Measure, 2013-2015*

Measure Type	Accounts	Gas Savings (m^3)
Building Shell	67	3,788,169
Controls	486	25,238,860
Greenhouse	17	5,295,971
Heat Recovery	23	1,696,982
Optimization and Maintenance	182	18,400,956
Other Equipment	107	36,085,459
Process Heat	12	4,179,649
Space Heat	765	39,517,947
Steam and Hot Water	233	4,076,918
Ventilation	211	7,139,452

*These data include duplicates of 2014 to represent the 2015 data.

Union Participant Data

The participant data files provided by Union included energy efficiency projects completed during the 2013 and 2014 program years. At the time of writing this plan, DNV GL does not have data for the 2015 program year. In the 2013 and 2014 Union tracking data, there are 803 records and 392 unique accounts. The records in the tracking data are per installed measure, so there may be multiple rows per project if more than one measure is installed, and there may also be more than one project per account. Across the three program years, we expect to have approximately1,300 records for approximately 600 unique Union accounts.

Table 6: Union Participation Metrics by Year

Program Year	Accounts	Gas Savings (m^3)
2013	197	109,243,796
2014	260	176,508,753
2015 projection*	260	176,508,753

*These data are duplicates of 2014 for reference purposes only.

We used the *project type*, *equipment type*, and *project category* variables in the tracking data to categorize measures. The resulting measure types are presented in Table 7. The largest measure types in terms of savings were maintenance, heat recovery, and building shell. The tracking data we received had some conflicting identifiers that made it difficult to create consistent measure types across the sample frame. For our initial sample design we divided the measures into two categories: equipment and operations & maintenance (O&M), but we plan to revisit this decision once we receive the complete dataset and we know the distribution of completed measures.

Table 7: Union Participation Metrics by Measure, 2013-2015

Measure Type	Accounts	Gas Savings (m^3)
Ag and Greenhouse	65	45,958,821
Building Shell	302	56,606,840
Controls	70	23,204,063
HVAC	59	39,719,864
Heat Recovery	89	71,423,260
Maintenance (O&M)	179	179,305,508
New Construction	17	3,815,481
Optimization (O&M)	62	27,153,170
Other Equipment	38	1,137,342
Steam and Hot Water	45	13,936,954

*These data include duplicates of 2014 to represent the 2015 data.

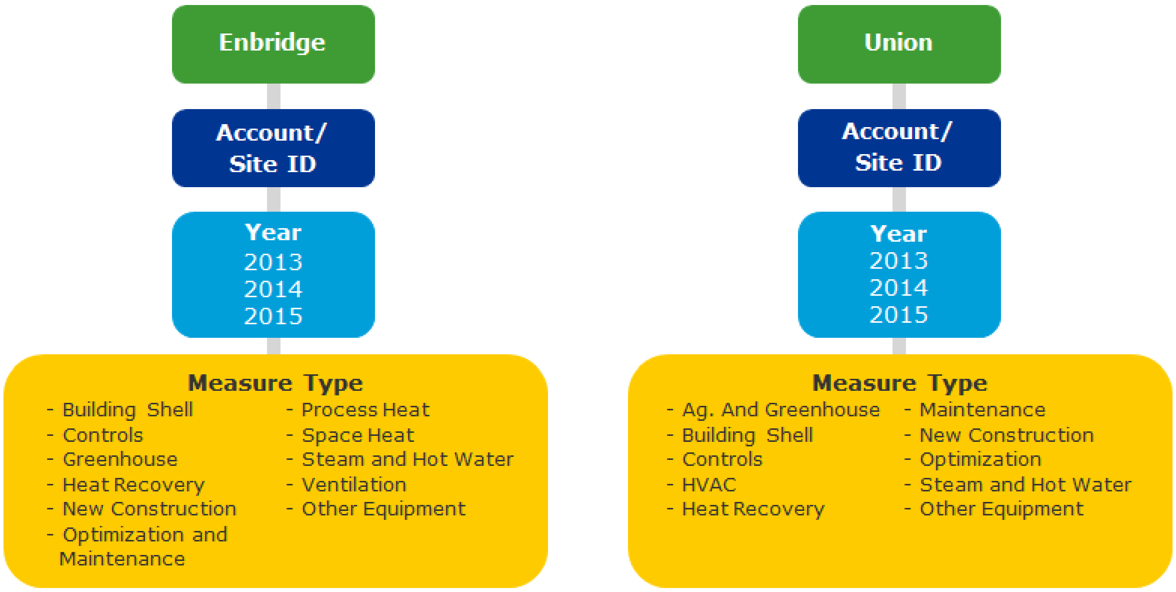
Task 2.2: Define the Unit of Analysis

Following data exploration, we defined the *unit of analysis*, which defines the level at which data will be analyzed, but not the level at which it will be collected, which is the *sampling unit*. We further discuss this distinction and how the sampling unit is defined in the Task 2.5 section.

The unit of analysis for this evaluation is a slight aggregation of the records in the tracking database. The purpose of the aggregation is to reduce the number of questions asked in each survey and to reduce the difficulty of parsing decision-making across multiple similar measures. We aggregated across elements that are likely to have less effect on decision making and did not aggregate across distinctions that are likely to play a larger role in how decisions were made.

The unit of analysis used in this sample design, presented in Figure 1, aggregates the data to the utility, account, year, and measure type. For Union, aggregating the tracking data to the unit of analysis reduced the number of records from 803 to 606 (2013 and 2014 data). For Enbridge, the number of records decreased from 1,603 to 1,471 (2013 and 2014 data).

Figure 1: Unit of Analysis



For this sample design, the unit of analysis and the sampling units are defined differently. While a unit of analysis separates units of different accounts/sites, program years or measure types, the sampling unit is specific to the customer. As an example, one Enbridge customer may have installed a new boiler in 2013 and insulation in 2014 which would be two different units of analysis, but since they were installed by the same customer they belong to one sampling unit. In the analysis phase, weights will be developed for each unit of analysis (account-measure type-year), but for the standard error calculation, data collected from a single customer (sample unit) will be a treated as a cluster rather than evaluated as if they are independent observations

Once aggregated to the unit of analysis, Union had an average of 1.5 units of analysis per account, while Enbridge has an average of 1.2 units per account across the 2013 and 2014 program years. In general, Union accounts tended to have more units of analysis per account than Enbridge accounts. Even so, with a handful of exceptions there were no more than 4 units of analysis per account. This will facilitate data collection, since it's generally reasonable to ask about 3-4 units, especially if 2 of them are the same measure in multiple years.

At this time we are unable to comment on the number of units per customer, though we know accounts some customers will likely have multiple accounts. Customer contact information will be requested in a documentation request following receipt of the final tracking data.

Task 2.3: Stratify the Data

As this is an initial pass at the stratification for a census, we followed a decision making process consistent with sample design, but knowing that we will be reviewing these decisions and making adjustments after data collection is complete. There is a balance between having too many and too few strata. In sample designs, more strata allow the design to control representativeness and estimated precision along more dimensions. Having more strata can hurt overall precision if variation across strata is less than or equal to variation within each stratum. Our initial stratification design has 108 strata defined by:

- Utility - We are treating each utility's programs as separate populations.
- Commercial vs Industrial vs Run-it-Right – decision making at the broad segment level is likely different due to the different business structures, whereas the design of the Run-it-Right program is very different from the other Enbridge offers.
- O&M vs Equipment – Decision making on equipment is more different from that on O&M than it is on types of equipment
- Program Year – Program year determines the data collected (free ridership vs. spillover or both)
- Size of unit (m^3) - when using ratio estimation with m^3 saved as a weight in the results, stratifying by size ensures that large measures affect the result proportionally and do not have an outsize effect. If this is not a census, stratifying by size would ensure large measures were included in the sample, which is important for the ultimate precision of the study.

Optimization and maintenance measures were grouped separately from the other measures, which involved installing or removing equipment.

Enbridge Stratification

The Enbridge stratification process is presented in Figure 2 and the tracking data summary by the groupings is presented in Table 8. Run-it-Right will follow the same process as the commercial and industrial segments.

Figure 2: Enbridge Stratification

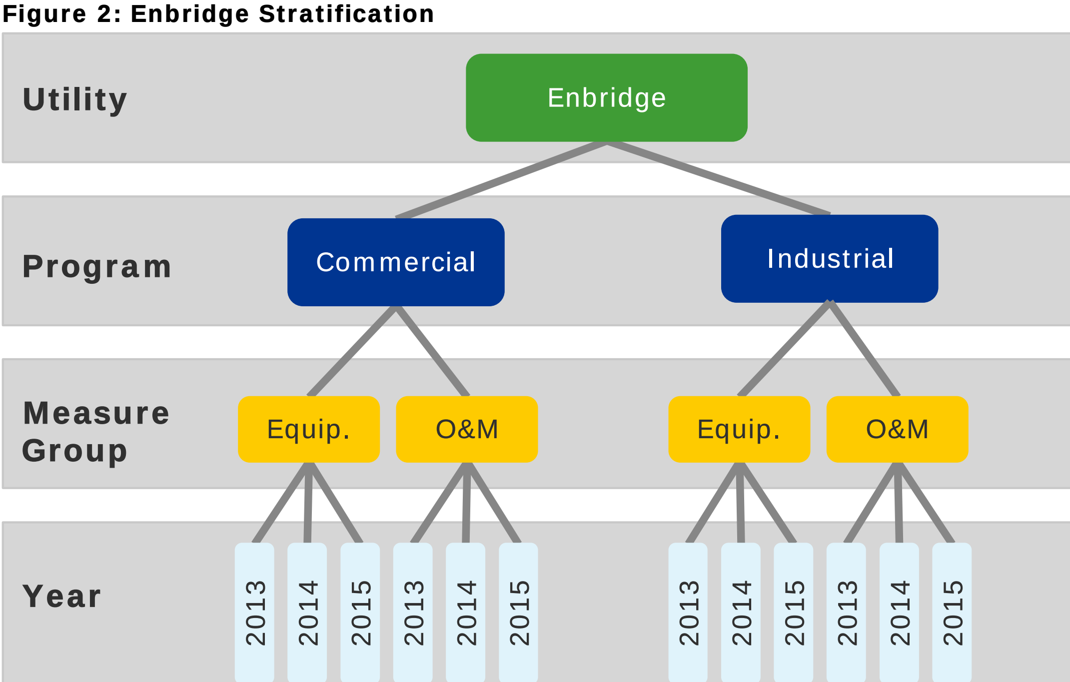


Table 8: Enbridge Participation Metrics by Grouping*

Utility	Program	Measure Group	Year	Accounts	Gas Savings (m^3)
Enbridge	Commercial	Equipment	2013	539	26,126,210
			2014	460	21,371,289
			2015	460	21,371,289
		O&M	2013	53	2,584,681
			2014	33	2,175,656
			2015	33	2,175,656
	Industrial	Equipment	2013	77	22,405,997
			2014	74	17,872,311
			2015	74	17,872,311
		O&M	2013	24	1,913,445
			2014	27	4,775,759
			2015	27	4,775,759

*These data are duplicates of 2014 for reference purposes only.

Union Stratification

The Union stratification process is presented in Figure 3 and the tracking data summary by the groupings is presented in Table 9.

Figure 3: Union Stratification

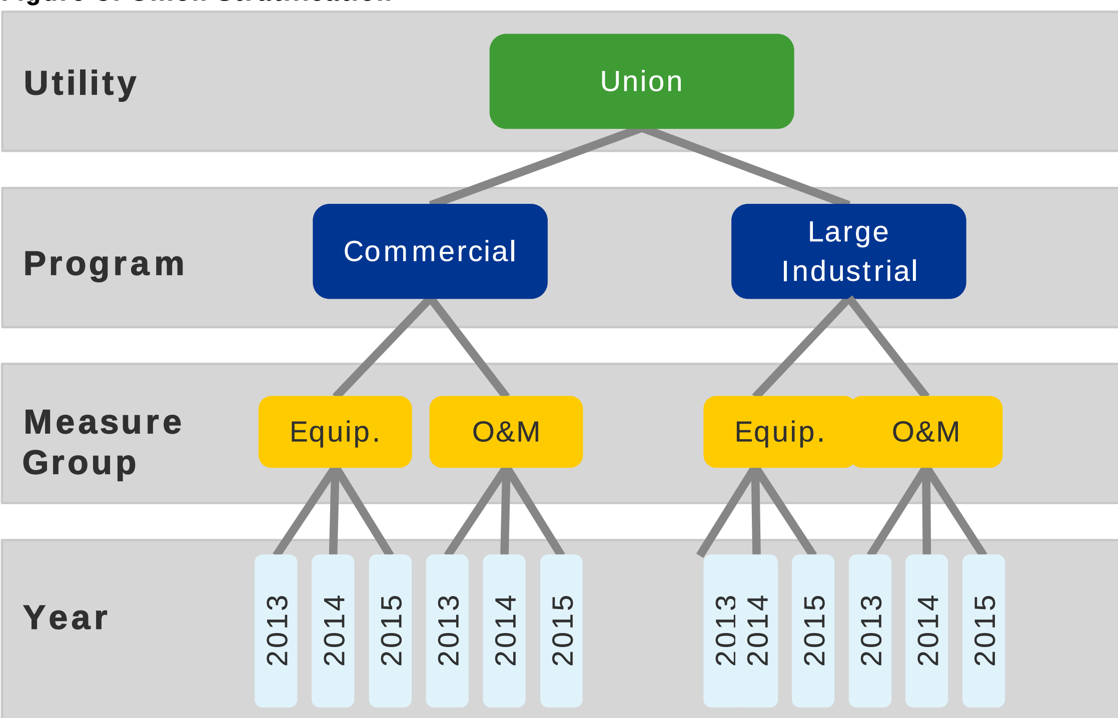


Table 9: Union Participation Metrics by Grouping

Utility	Program	Measure Group	Year	Accounts	Gas Savings (m^3)
Union	Commercial	Equipment	2013	147	28,658,112
			2014	184	45,508,018
			2015	184	45,508,018
		O&M	2013	38	12,823,518
			2014	58	17,855,569
			2015	58	17,855,569
	Large Industrial	Equipment	2013	21	50,632,883
			2014	28	42,747,797
			2015	28	42,747,797
		O&M	2013	16	17,129,283
			2014	36	70,397,369
			2015	36	70,397,369

*These data are duplicates of 2014 for reference purposes only.

Task 2.4: Design the Sample

Prior to completing the initial sample design, we assigned the data collection type for each customer.[9] We did this in order to determine the maximum expected number of respondents since IDI and CATI services have different response rates. We make these decisions at the customer level, rather than the unit of analysis (which includes measure group and year) to estimate the maximum expected number of respondents given the different expected response rates for IDI and CATI surveys. While the data collection will be completed at the sampling unit (customer), we assume that the account number provides a reasonable approximation.

We separated the preliminary IDI sample frame from the CATI sample frame based on three decision rules:

1. All Direct Access program measures
2. Customers with more than two measures
3. The largest customers up to a maximum IDI sample frame of 122 Union accounts and 90 Enbridge accounts. (106 IDIs with a 50 percent response rate).

In the final design, once we have project documentation complexity of measure will also be a consideration.

While we assumed a 50 percent response rate, our interviewers will be taking steps to ensure the highest response rate possible as detailed in Task 3: Data Collection. By assigning the data collection method at this point, we will be able to determine whether there are enough accounts assigned to CATI to use a sample design or if we should use a census. Figure 4 and Figure 5 present the number of units of analysis per account, which are the number of units about which a respondent would be asked, by data collection type using the 2013 and 2014 tracking data only.

[9] Currently this is at the account level, but will be at the customer level once we are able to identify customers with multiple accounts.

Figure 4: Enbridge Units of Analysis, 2013 and 2014

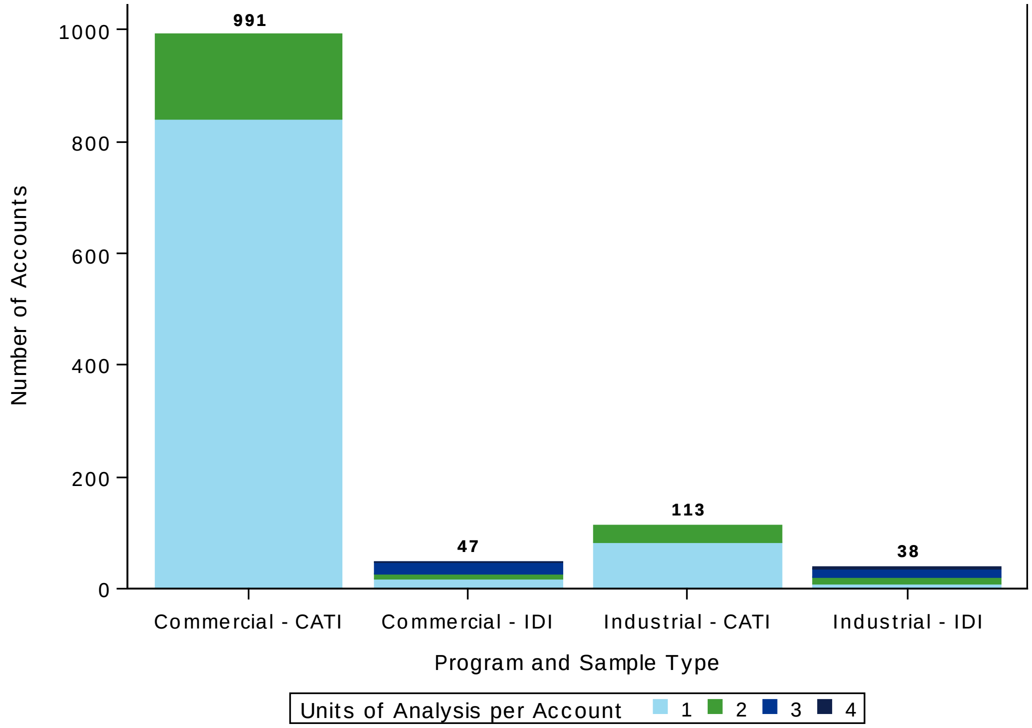


Figure 5: Union Units of Analysis, 2013 and 2014

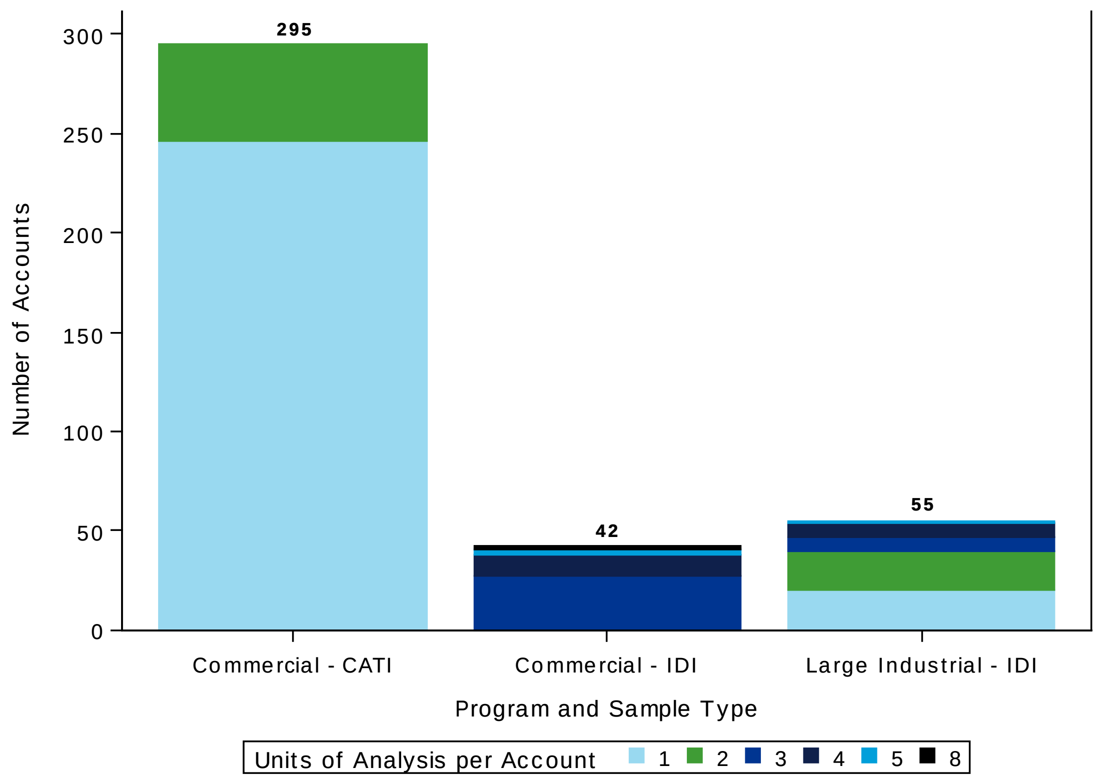


Table 10 summarizes the estimated sample sizes and the corresponding precision, for each overall program. Sample sizes are estimated based on our expected response rate attempting a census for all programs and years. Consistent with our recent experience in data collection for custom C&I programs, the sample sizes assume a 50 percent response rate for the IDI sample and a 33 percent response rate for the CATI sample. Since 2015 data is not currently available, we are using a copy of the 2014 program year as a placeholder for the 2015 data.

Table 10: Estimated Number of Completes and Sample Frame (Analysis Units)

Utility	Program	2013		2014		2015 Projected		All Three Years	
		n	N	n	N	n	N	n	N
Union	Comm and Small Ind	71	208	98	284	98	284	267	776
	Large Industrial	17	41	31	73	31	73	79	187
	Overall	**88**	**249**	**129**	**357**	**129**	**357**	**346**	**963**
Enbridge	Commercial	216	683	175	558	175	558	566	1,799
	Industrial	40	116	41	114	41	114	122	344
	Run-It-Right								
	Overall	**256**	**799**	**216**	**672**	**216**	**672**	**688**	**2,143**

Note; n= sample size (estimated number of completes), N= estimated number in sample frame

The study seeks to achieve 90/10 overall precision representing future participation. To project to future participants, treated as effectively an infinite population, we developed the precision estimate for the study without applying the Finite Population Correction. The table shows our estimates for the go forward non-FPC precision for each program. These estimates are based on the 33% CATI and 50% IDI response rate assumptions, a 0.7 error ratio (estimate of variance) the 2013/2014 program year data, and the stratification described above. Free ridership is based on 2014 and 2015 participants, while spillover is based on 2013 and 2014 participants. The final achieved precisions will depend primarily upon our response rates for the large customers.

Table 11: Anticipated Sample Sizes and Precision as Estimate for Future Program

Utility	Program	Free ridership		Spillover	
		n	RP	n	RP
Union	Comm and Small Ind	196	5%	169	5%
	Large Industrial	62	11%	48	17%
	Overall	**258**	**6%**	**217**	**10%**
Enbridge	Commercial	350	5%	391	5%
	Industrial	82	10%	81	10%
	Run-It-Right				
	Overall	**432**	**5%**	**472**	**5%**

Note; RP = relative precision at the 90% confidence level

Task 2.5: Prepare the Sample Frame

Once we have completed the final sample design, we will submit a data request to the utilities. The specific types of information we will be requesting are outlined in Table 12. The decision maker may not necessarily be located at the site where the project occurred and may be the same for multiple projects at multiple sites. The technical expert is someone who will be able to answer questions regarding the specific engineering specifications of the equipment. Vendors are the third party firms that were involved in the sale or design of the equipment, or the sale and performance of the O&M services.

Table 12: Information to Be Requested

Requested Information	Project Year		
	2013	2014	2015
Site Address	√	√	√
Project Documentation	√	√	√
Decision Maker Contact Information: ▪ Full Name ▪ Role ▪ Mailing Address ▪ Email Address ▪ Direct Business Phone Number	√	√	√
Technical Expert Contact Information: ▪ Full Name ▪ Role ▪ Mailing Address ▪ Email Address ▪ Direct Business Phone Number	√	√	
Vendor Contact Information: ▪ Full Name ▪ Role ▪ Mailing Address ▪ Email Address ▪ Direct Business Phone Number	√	√	√

Once we have received the requested contact information, we will identify instances where a contact was involved in multiple projects, even across sites. While the projects are conducted at the site level, the decision maker, technical expert, or vendor may have been involved in projects at multiple sites. For example, multiple participating sites for the same retail chain may have one energy manager from the corporate office but the technical expert may be site specific. Using this contact information and taking into account cross-site involvement, we will assemble the CATI and the IDI sample frame.

Task 3: Data Collection

Table 13: Task 3 Subtasks and Completion Status

Task 3: Data Collection
☒ **Task 3.1: Program Managers and Staff Interviews**
☒ Union Gas Program Portfolio Management
☒ Enbridge Commercial Program Managers
☒ Enbridge Industrial Program Managers
☒ Enbridge Run-It-Right Program Managers
☐ **Task 3.2: Program Energy Solutions Consultants Interviews**
☐ Energy Solution Consultant Interview Guide
☐ 10 pre-survey interviews
☐ 10 follow up interviews
☐ **Task 3.3: Program Participant CATI Survey**
☐ CATI survey Instrument
☐ CATI survey interviews completions attempted
☐ **Task 3.4: Program Participant In-Depth Interviews**
☐ In-Depth Interview Instrument, mirroring CATI instrument
☐ IDI completions attempted
☐ **Task 3.5: Program Participant Engineer Interviews**
☐ **Task 3.6: Participating Vendor In-Depth Interviews**
☐ In-Depth Interview Instrument
☐ 80 IDI completions attempted

Note; the number of CATI and IDI completion attempts will be filled in once we receive the final dataset.

Objectives

To inform this NTG evaluation, the evaluation team will collect information from both Custom C&I program participants and key program actors. The following table shows the participants and key program actors we plan to target and what aspects of influence on the energy efficient project we are planning to explore. We expect that for some larger energy efficiency projects, but not all projects, the Energy Solutions Consultants will be familiar with some project-specific details.

Table 14: Aspects of Influence on the Energy Efficient Project

Aspects of Influence	Program Participants	Participating Vendors	Program Managers	Program Energy Advisers
Program influence on the participant's general practices	√		√	√
Program influence on the vendor's general practices and equipment recommendations		√	√	√
Program influence on the specific project	√	√		√ *
Vendor influence on the specific project	√	√		√ *

*(possibly for larger projects)

Activities

To inform this NTG evaluation, the evaluation team will collect information from both Custom C&I program participants and key program actors. The following table shows the participants and key program actors we plan to target and what aspects of influence on the energy efficient project we are planning to explore. We expect that for some larger energy efficiency projects, but not all projects, the energy advisers or utility account representatives will be familiar with some project-specific details.

The following is a summary of the number of interviews we plan to complete. A more detailed breakdown of our target number of surveys and interviews is provided in the description of the methodology in Task 2: Sample Design. As discussed in that section, many of the estimates of the targeted number of completed surveys are preliminary pending more precise information concerning the size of the participating customer populations, the mix of customer sizes, the mix of project sizes, the size of the participating vendor population, the number of participant spillover projects, etc.

Table 15: Target Number of Completed Surveys/Interviews

Phase 1	Phase 2			Phase 3	
Program Managers/Staff (In-Depth Interviews)	Program Energy Advisers[10]	Program Participants (CATI Surveys)	Program Participants (In-Depth Interviews)	Program Participants (Engineering Follow-Up Interviews for Spillover)	Participating Vendors (In-Depth Interviews)
Up to 6	10 + 10	≤2,200 *	≤430 *	≤80 *	≤80 *

*Note; The number of CATI and IDI completion attempts will be filled in once we receive the final dataset. All numbers represent the maximum number of surveys or interviews.

Shortly after the scope of work is finalized and approved, DNV GL will submit draft interview guides and CATI surveys to the EAC for review. The final interview guides and survey instruments will address any comments or suggestions from these reviewers. While this review process is ongoing, we will also check with the EAC and the relevant utility and program contacts to insure that we are talking to the appropriate people, have the necessary contact information, and have an advance letters from the utilities.

Timing

DNV GL recognizes the limitations of the calendar in conducting survey research. During summer months, holidays increase the difficulty in reaching individuals. DNV GL will take efforts to conduct the majority of data collection before the height of summer holidays. Typical survey protocol dictates that contact with a survey respondent should be attempted 6-8 times before being considered ‘exhausted’; DNV GL will adapt survey protocols to ensure that contact with an individual is not attempted more than 2x in a given calendar week and 3x in any two weeks to ensure that holidays do not influence response rates.

Informed Respondent

For data collection efforts involving non-program staff (e.g., participant surveys, participant interviews, participant follow-up interviews, participating vendor interviews), DNV GL will include a question battery designed to ensure that only informed respondents are participating. For participating customer

[10] Program Energy Advisors will be interviewed both in Phase 1 and Phase 2.

respondents, DNV GL will define informed respondents as interviewees who directly participated in the project(s) in question. For participating customer respondents, DNV GL will make every effort to reach informed respondents. We define informed respondents as interviewees who were:

1) Aware of the program at the time of the project; and
2) Either directly involved in the decision to choose equipment and go forward with the project or reasonably familiar with the project's decision-making process.

Some companies with multiple projects and diverse decision makers may require multiple interviews. We will not administer survey for projects where the informed respondents are not available. DNV GL will include a battery in each relevant instrument aimed at confirming the individual interviewed is an appropriate informed respondent.

Response Rates

Survey response rates have been in decline over the past decade. This is especially true for residential surveys, where cell-phone only households have made surveying difficult, but there has also been erosion of response rates for business surveys. In order to achieve increased response rates, DNV GL will prompt program participants with both advance emails and advance letters, informing them of the survey and requesting participation. Advance letters, sent through traditional postal mail, are generally better received (and read) when sent by the recognized energy provider and should be sent on utility letterhead, if possible.

In order to execute the mailings, it is critical that DNV GL be provided with accurate contact information for the correct informed respondent. This will include, but is not limited to, the correct individual's:

- Full Name
- Role
- Mailing Address
- Email Address
- Direct Business Phone Number

DNV GL will send the above-mentioned emails and letters to all program participants included in the CATI sample frame as well as those program participants identified for IDI, and participating vendors. For IDIs, there is an additional opportunity to improve response rates – providing respondents with the opportunity to schedule their own interview time. DNV GL will accomplish this with either an invitation to email DNV GL directly about preferred times or will utilize an online scheduling service where individuals may choose their own preferred times.

Handling of Optimization and Maintenance Projects

Optimization and Maintenance projects will be separated from equipment installation in the sample design and require special consideration for data collection as well. Maintenance projects in particular are by their nature recurring. The question of how to credit the program for maintenance this year when the customer participated in the past is complex. DNV GL and the TEC considered this issue while finalizing the contract and decided that the primary objective of the free ridership estimation will be to capture the effect of the program(s) on the current project. The effect on the current project of prior and indirect program experience will be captured in a secondary, less rigorous question sequence.

The primary attribution questions will be framed by questions that ask about decision making for the current project alone so that the scored attribution sequence will capture the effect of the program on the current project. After the scored section of the survey is complete we will capture the indirect, longer term attribution effect by asking:

- "Now, without *any* utility assistance for *any* projects in the past, on a scale of 1-10 what is the likelihood that you would have <taken this EE Action>?

The maximum of the primary attribution and this score will provide us with an idea of how much higher attribution would be if a longer term view were taken.

To limit customer burden and ensure the validity of our spillover analysis we will limit the investigation of secondary attribution to:

- *Measures with less than 100 percent primary attribution*: if primary attribution is 100 percent, then secondary attribution is as well
- *2015 participants*: 2015 participants will not be asked spillover questions. It would be awkward to ask about spillover and then secondary attribution. It could affect the results to ask about secondary attribution prior to asking about spillover.

Deliverables

- Program participant CATI survey instrument (draft and final)
- Program participant IDI guide (draft and final)
- Participating vendor IDI guide (draft and final)
- Program manager and staff IDI guide (draft and final)
- Program Energy Solutions Consultant IDI guide (draft and final)
- CATI and IDI participation email & mail scripts

Task 3.1: Program Managers/Staff (In-Depth Interviews, Phase 1)

In order to better understand program logic, methods, execution, and intent, DNV GL conducted IDIs with program managers and then program staff. These interviews focused one:

- Details of how the program design
- Details of how the program is implemented and marketed
- Understanding the program theory and logic
- Identifying key staff such as Energy Solutions Consultants and what roles they play,
- Identifying how decisions are made
- Identifying how communication between parties occurs.

DNV GL staff interviewed program staff from Enbridge and Union on the following dates:

Table 16: Program Manager Interviews

Date	Company	Program
1/22/2016	Union Gas	Program Management - Portfolio
1/25/2016	Enbridge	Commercial Programs Interview
1/29/2016	Enbridge	Industrial Programs Interview
1/29/2016	Enbridge	Run-It-Right Program Interview

Task 3.2: Program Energy Solutions Consultants (Phases 1&2)

Ten Energy Solutions Consultants will be interviewed by experienced DNV GL staff prior to the creation of program participant survey instruments, in order to better inform those instruments. Five ESCs will be interviewed from both Enbridge and Union Gas (10 total interviews), with a distribution of consultants speaking to all commercial and industrial programs. An original list of topics has been modified following interviews with program managers. Topics for interview will include:

- Their typical responsibilities
- The nature of their routine communications and interactions with Custom C&I customers and how this might vary with the size of the customer or the customer type (e.g. chain stores)
- How they target program recruitment at Custom C&I customers of certain types or in certain areas
- Nature of program recruitment; communication type by customer size and rate class
- How they become aware of potential energy efficiency projects
- How they promote energy efficiency
- How they identify which customer representatives are the key project decision makers
- Which customer representatives are the key project decision makers and how this might vary depending on company/organization size or type
- At what stage in project development they typically get involved with a project
- How many projects are typically rejected
- What are the barriers to program participation and how they try to mitigate them
- What information, financial incentives or technical assistance they offer to Custom C&I customers for energy efficiency projects
- What they perceive to be their most valuable contributions to the development of energy efficiency projects
- How frequently the rely on program technical support staff for project support
- How closely they monitor the progress of active projects
    - If there is any evidence of project cancellations due to free ridership
- Whether they have received any training or guidance on how to minimize free ridership
- Whether there are any warning signs that a project might be a free rider
- What roles trade allies play in project identification and how they interact with them
    - How trade allies work to mitigate free ridership

Following initial interviews and surveys of program participants, DNV GL staff and engineers will return to call up to ten of Energy Solutions Consultants to collect additional information about specific projects that will be useful for measuring program attribution. In such cases staff will ask project-specific questions such as:

- Whether they were involved in originating the project idea and, if so, what was their role
- Whether they were involved in planning and the development of the project details, and if so, what was their role

- Whether they were involved in the decision to go ahead with the project, and if so, what was their role
- At what stage in this project conception, planning and development process they got involved
- Whether they thought the availability of the Custom C&I program financial incentives had any influence on the timing or character of the project and if so, what was the nature of this influence
- Whether they thought any other Custom C&I program services (e.g., training, audits, technical assistance, helping find a vendor, selling the project to upper management, etc.) that the participant received had any influence on the timing or character of the project and if so, what was the nature of this influence

Task 3.3: Program Participant CATI Survey (Phase 2)

For the CATI surveys and IDIs with participants, we are developing flexible instruments that will have different modules depending on the year in which the Custom C&I customer participated. Table 17 shows how these modules will be distributed across the program years.

Table 17: Net-to-Gross Survey Modules Depending on Program Year

Net-to-Gross Modules	2013 Participants	2014 Participants	2015 Participants
Free Ridership	√*	√	√
Spillover	√	√	

*The free ridership estimates from this program year will only be used to inform the spillover analysis and will not be used for the core free ridership analysis.

There is no spillover module for the 2015 participants because we are assuming that not enough time has elapsed for the large majority of these participants to have done a spillover project. It is possible that some of the largest customers may have done a spillover project in this short timeframe. Since these large customers (3 or more projects) will be covered by IDIs, we will give the interviewers the flexibility to administer the spillover questions if a spillover project is identified. However, our survey cost estimates assume that for most of the 2015 participants we will not administer the battery of spillover-related questions. As the table indicates, we will be collecting free ridership information from the 2013 participants because this information is required for our participant spillover methodology, but we will not be using these data for our core free ridership calculations because we would prefer to use more recent program years.

OUTLINE:

- **Introduction**
- **Informed Respondent**
    - o Cite specific project, determine involvement
        - Program awareness
        - Equipment choice
    - o Role
    - o Responsibilities
    - o Ask about how long at company (since before the project date?)
- Identify names of other primary project contacts, for potential follow up conversation with DNV GL engineer

- **Organization Policies and Purchase Procedures**

- **General Program Awareness and Interactions**
- **Specific project verification (Framing)**
  - When first considered?
  - Reasons for project?
  - Major sources of info?
  - The general decision-making concerning energy related purchases and practices
  - Who in their company or organization makes decisions about equipment replacement and retrofit projects and how this might vary with the size or cost of the project
  - What information sources are used in making these decisions
  - Whether the company/organization has any formal requirements or informal guidelines about the purchasing of energy using equipment and, if so, what are these requirements/guidelines
  - Whether their company has a corporate "green " mandate
  - The development of the specific program-incentivized project
  - Where the idea for the project originated and who were the key persons involved in the project conception -- whether within the participant's company/organization or without (e.g., vendors, Custom C&I program Energy Solutions Consultants)
  - Who was involved in the planning and development of the project details
  - Who was involved in the decision to go ahead with the project
  - At what stage in this project conception, planning and development process did the Custom C&I program get involved
  - Whether the program provided any services to the respondent's company/organization beyond the financial incentives (e.g., training, audits, technical assistance, helping find a vendor, selling the project to upper management, etc.). To inform the free ridership questions, the interviewers will have information on all program activities reported by the tracking databases, but this query is designed to collect information on program activities that may be unreported and also to find out which program activities were top-of-mind for the respondent.
  - Whether the project changed from its original conception and what these changes were and why they were made

- **Direct attribution battery**
  - Determining the overall influence of the program, along with program effects on
    - Timing
    - Efficiency
    - Sizing or Quantity

- **Spillover battery**
  - Inquire about additional projects *after* other projects[11]
    - First check to ensure not rebated
    - Project type
    - Project data
    - Project contact

[11] Information collected regarding additional projects will be used not only to calculate spillover, but to check against program records and ensure that the project was not a tracked project with direct attribution.

- Project location
- Project dates

- **Firmographics**
  - Business type and
  - Business size (ft2 and # of employees)
  - whether they lease or own their facilities

- **Closing**

Task 3.4: Program Participant In-Depth Interview (Phase 2)

In addition to executing telephone interview surveys for standard projects, DNV GL staff will conduct IDIs with participants of particularly large or complex projects. The subjects covered in the IDIs are the same as with standard projects, following the same outline. In DNV GL's experience large and complex projects do not lend to pre-programmed interviews in the same way that standard (single) projects do, as it may be necessary to speak with multiple individuals or to dive deeper into questions to determine answers to questions than can be completed in a fully programmed interview. Conducting IDIs of customers with large or complex projects is a standard method for DNV GL, with experienced and expert interviewers conducting all interviews. These interviews are typically conducted with the 'decision maker' – an informed respondent who has at least some say in whether or not to proceed with a project and is aware of the project's impacts.

Task 3.5: Program Participants (Engineering Follow-Up Interviews for Spillover, Phase 3)

For some projects, it will be necessary to follow up with an additional individual or individuals, aside from the 'decision maker'. Engineering follow up calls are a specialized form of IDI that are conducted between a DNV GL engineer and an individual at the customer site that can speak to the specific engineering specifications of the equipment. DNV GL will ask specific questions that will allow for the calculation of energy savings.

These interviews will be individually tailored, depending on equipment installations, with the goal of gaining information to calculate energy savings.

Task 3.6: Participating Vendors (In-Depth Interviews, Phase 3)

Vendors that performed work on projects identified in the sample will also be interviewed. IDIs with up to 80 of these vendors will follow the following outline:

- **Introduction**
- **Informed Respondent**
- **Company background**
  - Which products or services they sell
  - Which types of C&I customers they typically do business with
  - What the size of their company is

- **Sales and marketing**
  - How they promote energy efficiency
  - How they identify which customer representatives are the key project decision makers
  - Which customer representatives are the key project decision makers and how this might vary depending on company/organization size or type
  - What role the Custom C&I program incentives play in their sales pitches

- **General program involvement and influence**
  - How they became involved with the Custom C&I program
  - Why they became involved with the program
  - How frequently they offer program incentives
  - How frequently they interact with program staff
  - How they keep track of Custom C&I program incentives and requirements
  - Whether the Custom C&I programs have provided them with any sales leads
  - Whether they have received any training from the program
  - Whether there are other services that the program provide them
  - To what degree the Custom C&I program incentives and other services influence the implementation of energy efficiency projects in the C&I sector
  - What types of C&I customers are more likely to be influenced by program incentives and which types are less likely to be influenced
  - Whether they are offering energy efficient products or services through the program that they did not offer before becoming involved with the program
  - Whether they are recommending energy efficient products or services more frequently now than they did before becoming involved with the program
  - Whether they have suggestions as to what kind of interventions would increase the program's influence in the project

We will call back some of the vendors to collect additional information about the project that will be useful for measuring program attribution. In such cases we will ask some project-specific questions such as:

- Whether they were involved in originating the project idea and, if so, what was their role (informed respondent)
- Whether they were involved in planning and the development of the project details, and if so, what was their role
- Whether they were involved in the decision to go ahead with the project, and if so, what was their role
- At what stage in this project conception, planning and development process they got involved
- Whether they thought the availability of the Custom C&I program financial incentives had any influence on the timing or character of the project and if so, what was the nature of this influence
- Whether they thought any other Custom C&I program services (e.g., training, audits, technical assistance, helping find a vendor, selling the project to upper management, etc.) that the participant received had any influence on the timing or character of the project and if so, what was the nature of this influence.

Task 4: Data Analysis

Table 18: Task 4 Subtasks and Completion Status

Task 4: Data Analysis Subtasks
☐ **Task 4.1: Analyze Survey and Interview Data**
☐ Sample frame data transformation
☐ Sampling weight
☐ Data validity check
☐ **Task 4.2 Calculate estimates**
☐ Attribution
☐ Spillover
☐ Net-to-gross

Objectives

Once a critical mass of CATI surveys and IDIs have been completed, we will begin the analysis phase (in parallel with the completion of data collection). In this section, we describe the initial survey and interview data analysis as well as the calculation of attribution, spillover, and NTG.

Task 4.1: Analyze survey and interview data

The analysis flow after data collection begins with transforming the collected data back to the level of the unit of analysis. This translation depends on the number and grouping of program measures or projects asked about for an individual customer, and whether subsampling was required.[12]

The survey will collect attribution information on each measure type. We apply the free ridership and spillover "scoring" methods to determine the free rider and spillover factors for each measure type. We then apply these factors to the associated gross savings to produce net-of-free riders and spillover savings for each measure type. Data collected from a single customer will be treated as a single cluster in error estimates.

We will use the sampling weights created during the sample design process to expand the customer sample in each sampling cell (stratum) to represent the full participant population in that cell. Targeted cells for which we are unable to obtain any responses will either be treated as not represented by the sample, or will be collapsed with other cells for sample expansion.

The application of attribution and spillover algorithms that convert survey and interview data into energy savings values ready for expansion involves consistency checks for each respondent. These checks utilize both questions directly used in the algorithms and verbatim responses that contain information on the reasoning of the respondent's responses.

Task 4.2: Calculate Estimates

The last analysis step involves expanding the attribution and spillover savings to the population via ratio estimation, and generating the combined NTG ratio for each segment of interest. . We will estimate inside

[12] Based on the initial data we do not anticipate needing to subsample: few accounts had more than two units. This may change once we have 2015 data and information that allows us to identify unique customers.

and outside spillover (inside spillover occurs at the same site as the program measure) separately for each segment, and sum them to determine total spillover.

We will then calculate corresponding ratios across the segment:

Equation 1: Free Rider Savings

$$FR = \frac{Weighted\ sum\ of\ NFR}{Weighted\ sum\ of\ G}$$

Where:

NFR = Net of free rider savings

G = Gross savings

Equation 2: Net of Free Rider Savings

$$NFR = 1 - FR$$

Equation 3: Inside Spillover

$$SO_{inside} = \frac{Weighted\ sum\ of\ S_{OI}}{Weighted\ sum\ of\ NFR}$$

Equation 4: Outside Spillover

$$SO_{outside} = \frac{(Weighted\ sum\ of\ S_{OO})}{Weighted\ sum\ of\ NFR}$$

Equation 5: Total Spillover

$$SO = SO_{inside} + SO_{outside}$$

Where:

S_{OI} = Inside spillover savings (0 for customers with no spillover)

S_{OO} = Outside spillover savings (0 for customers with no outside spillover)

Equation 6: Net-to-gross Ratio

$$NTG = (1 - FR)(1 + SO)$$

We calculate spillover as a fraction of net of free rider savings, but this can also be reported as a fraction of gross savings.

Task 5: Reporting

Table 19: Task 5 Subtasks and Completion Status

Task 5: Reporting Subtasks
☐ **Task 5.1: Monthly Status Reports**
☐ **Task 5.2: Bi-Monthly Updates**
☐ **Task 5.3: Methodology Memo**
☐ **Task 5.4: Draft report**
☐ Include estimates of free ridership
☐ Include estimates of participant spillover
☐ Include forward free ridership and spillover data
☐ **Task 5.5: Final report and presentation**
☐ Final report addressing comments on draft report
☐ In-person presentation

Objectives

DNV GL recognizes the importance of providing clear and timely reports on project progress, evaluation activities and results.

Activities

Our approach to reporting for this project includes:

- **Monthly emailed status reports:** Every month the DNV GL project manager will submit a monthly status report to the EAC, via email, which will summarize the past month's activities, notify them of the next month's activities, and report on how closely the evaluation is adhering to the original schedule. However, if there are methodological questions or delays in responses to data requests that could put the evaluation off schedule, the program manager will notify the EAC of these issues immediately for proposed resolution so that the evaluation schedule is not compromised.

- **Bi-monthly study updates to the EAC:** the DNV GL project manager will provide the EAC with study updates via teleconference on a bi-monthly basis in alignment with scheduled EAC meetings. These bi-monthly study updates will provide similar information as in the monthly emailed status reports, although the more interactive format of the teleconference should allow for greater discussion and quicker resolution of any key issues.
- **Methodology Memo:** DNV GL will produce a memo detailing the methodology and rationale for the calculations to be used in the analysis. This memo will constitute most of the methodology section in the draft report and will allow the EAC to review and comment on the methods prior to receiving the results of the analysis.
- **Draft report:** At the conclusion of the evaluation, DNV GL will submit to the EAC a draft report that will present all the information requested in the RFP's research objectives including:
    - Estimates of program free ridership factors by market sectors and precision targets for both the Enbridge and Union Gas' custom C&I programs;
    - Estimates of participant inside and outside spillover[13] by market sectors and precision targets for both the Enbridge and Union Gas' custom C&I programs;
    - Guidance on the development of a strategy for applying free ridership and spillover data collected on previous program participation to forward looking DSM program activity.

[13] Potential electric spillover will not be reported in kWh, but descriptively, as electric spillover is outside the specific scope of this evaluation.

Along with these key findings, we will also show how these estimates were derived and what data from the IDIs and CATI survey were used to inform these estimates, including any qualitative findings regarding non-incentive based utility services.

- **Final report and presentation:** After receiving comments on the draft report from the EAC members, DNV GL will produce a final report which addresses all these comments along with a comment matrix that shows how we addressed them and why. We also plan to deliver an in-person presentation of these results to the EAC and the Ontario gas utilities.

Our current schedule has the project completion as November 16.

Table 22. Schedule of Deliverables

Tasks	2016											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Initial Staff Interviews	✓											
Kickoff Meeting			10									
Tracking Data Merge/Cleaning												
Draft Work Plan			2									
Work Plan Comments			17									
Final Work Plan			31									
Draft Survey and IDI Instruments			17									
Survey and IDI Instrument Comments			31									
Final Survey and IDI Instruments				14								
Review Project Documents												
Data Collection												
Data Analysis												
Methods Memo							29					
Draft Report									30			
Draft Comments												
Final Report & Presentation											16	

RISKS AND CONTINGENCIES

The tables in this section document the risks to project schedule, finances and quality and the contingencies DNV GL has in place to handle them.

Table 23: Schedule Risks

Schedule Risks	Explanation	Contingency
Data Reception	Timing Controlled by Union/Enbridge. Currently have no data for Run-it-Right and no data for 2015 program year. Anticipate data in March, but unsure of exact timing.	1) Move forward with work plan without missing data. Use 2014 data to proxy 2015.
Data Processing	Categorizing and Aggregating data in a way that is meaningful to eventual respondent and in the context of the NTG analysis. Initial data has some conflicting and unclear information.	1) Use best judgment with initial data for work plan. 2) Maintain list of "judgment calls" to clarify 2) confirm and clarify based on project documentation when it arrives 3) request clarifications where project documentation does not resolve questions.
Documentation Reception	Timing Controlled by Union/Enbridge	1) Send formal documentation request with explicit, agreed upon deadline for documents needed.

Schedule Risks	Explanation	Contingency
Documentation Reviews	Projects may have considerable documentation, some of which does not have bearing on the NTG effort.	1) Establish clear guidelines for information to be reviewed 2) provide full context of evaluation, goals and information needed from project documentation to satisfy project objectives 3) Utilize engineers familiar with a) custom gas projects and b) NTG methods and interviews
Contact information Reception	Timing Controlled by Union/Enbridge	1) send worksheet for contact information request 2) include clear directions for worksheet completion, including context of what we are attempting to learn from the interviews. 3) ensure the worksheet is simple and easy to complete.
Contact information processing	Contact information may be incomplete or come in hard to use format	1) Clear directions for the request 2) use experienced analyst to prepare data for survey
Resourcing	Having the right resources available at the right time is a challenge with projects that have experienced delays	1) reserve necessary resources for project in DNV GL's internal systems. 2) keep project on schedule to avoid conflicts with other project needs. 3) keep project sponsor aware of needs and championing project
Survey House Availability	Availability at the right time is a challenge with projects that have experienced delays	1) start conversation with Nielson now. 2) identify potential backups (Malatest) in case Nielson is unable to work on project due to long delay.
Review Periods	Dependent upon OEB/EAC/TEC priorities	1) establish clear and explicit deadlines for reviews
Decision Making	Dependent upon OEB/EAC/TEC priorities	1) schedule meetings with clear agendas that have key decisions up front.
Response Rates	Response rates on surveys have been declining, which can extend the time required for data collection	1) *attempt* a census so that call order does not matter, provided number of sample units makes this manageable and within original scope size. 2) IDI rather than CATI for the most complex and large projects 3) email participants prior to call to ask for cooperation 4) send advance letter to participants prior to call to ask for cooperation

Table 24: Financial Risks

Financial Risks	Explanation	Contingency
Currency Exchange Rates	USD/CAD rates have been highly variable	Fix prices in USD.
IDI Sample Size	Attempting a census with best practice approaches (minimum number of calls for all sites, and all efforts made to increase response rates) and an assumed 50 percent response rate could mean more completed IDIs than we have budgeted for.	1) Clear concise instruments to reduce time on phone 2) Simple data entry forms to reduce time entering data 3) Use admin staff to enter data for engineers who use paper form on phone (engineer review digital data after entry) 4) Advance letters and emails to decrease number of attempts per complete 5) limited number of dedicated engineers to reduce training costs and increase efficiency on attempted calls
Survey House re-negotiation	Project delays may prompt survey house to request re-negotiation	1) exchange rate may allow increase in survey costs if survey house paid in CAD. 2) open negotiations with alternative provider 3) census may provide survey house with cost efficiencies
Resourcing	Planned resources have been promoted since project scoped and now cost more. Planned resources have left company	1) substitute with acceptable alternatives. Inform TEC of new staff qualifications.
Travel	Additional kickoff meeting was not in original scope	1) cost may be absorbed by having OEB as decision make rather than consensus based TEC.
Timeline	Longer timeline tends to use project funds more than shorter timeline	1) seek to reduce schedule delays 2) Ensure efficiency or delay non-critical work when critical path is delayed to avoid additional expense

Table 25: Quality Risks

Quality Risks	Explanation	Contingency
Response Rates	Response rates on surveys have been declining, which can reduce sample sizes, introduce uncertainty about bias and make it hard to get data from large customers who have a large effect on final result	1) attempt a census so that call order does not matter. 2) IDI rather than CATI for the most complex and large projects 3) email participants prior to call to ask for cooperation 4) send advance letter to participants prior to call to ask for cooperation
Informed Respondents	Multiple people in a business are often involved in the decision to purchase capital equipment or spend money on optimizing or maintaining existing equipment. For consistency and cost reasons a single respondent from a company is preferable to interviewing multiple people at a business about the decision. Ensuring we have a respondent who knows enough about the decision to complete the project and the influence of the program on that decision is the crucial challenge of the data collection effort	1) Clear guidelines and screening questions to determine an informed respondent 2) removal from study of un-informed respondents 3) single interview for a project may require contacting multiple people at the site to determine an informed respondent.
Engineering Estimates	Spillover estimates will be based on engineer estimates of savings for projects that were not part of a program. We expect that these projects will not have the typical amount of documentation that we see for program rebated projects. The engineering estimates will be based on respondent provided information, and in some cases may not include specific sizes or operating characteristics.	1) Engineers will be required to thoroughly document information collected from the respondents and from third party sources. Justification for savings estimates will be provided, along with values and sources of key assumptions and calculation methods. 2) A senior engineer (Tammy) will review all estimates. 3) transparently provide documentation of project savings (within confidentiality limits) in appendix of report
Analyst Experience	Since the project start, turnover in our analyst group has meant a resourcing change for this project. The planned analyst has less experience, but has proven herself capable on similar projects	1) direct oversight of analyst by PM who has performed the same work previously 2) Any further substitution will be with an analyst PM has personal experience of success working with in the past.

APPENDIX A SAMPLING PROCESS

A sample is a collection of data items such as those collected through surveys, metering or onsite observation. A sample design is required when a sample does not include the entire target population. Most sample designs are driven by cost constraints (including schedule constraints), desired precision or both. The sampling process described here ensures that all bases are covered, ensuring optimal precision around estimates of interest for the data collected. The process we followed was:

1. **Identify Goals, Methods and Constraints**: for sampling, the goals consist of identifying the primary and secondary estimates of interest: what quantitative results are most important. Defining the data collection methodology –the process used to gather the data for the analysis – and the estimation method – the approach used to calculate the primary estimate of interest – is critical for defining elements of the design. Cost and schedule constraints surrounding the data collection and analysis then determine an upper bound for the sample size.

   - **Goals**: For this study the primary estimate of interest is the NTG ratio for each program. The NTG ratio is the parameter that we are targeting for 90/10 precision for each program.

     As will be described later in the methodology memo, we calculate the NTG ratio as

     NTG = (1-FR)(1+SO).

     Since spillover tends to be small, this formulation is mathematically very close to the simpler formula indicated in the recent Ontario evaluations

     NTG = 1-FR + SO.

     We prefer the multiplicative formula as a more complete expression of the relationship between free ridership and spillover.

     Previous work in Ontario indicates that free ridership is on the order of 10% to 60% across program segments, 50% overall on a savings-weighted basis. Spillover is on the order of 5%. Because spillover is generally small, the precision of the full NTG will in most cases be close to that of the net-of-free rider factor, even with a modest spillover sample size.

   - **Methods and Constraints:** We are using two data collection methods, each of which have different costs associated. Due to cost constraints we must limit our use of in depth interviews to those projects where it will make the most difference in the estimate: we have budgeted for 132 IDIs and these will be deployed on the largest and most complex projects as identified based on tracking data descriptions. CATI surveys will be used to collect the balance of the data that we do not have the funds to collect with IDIs. For smaller and simpler projects where the decisions made are more straight forward, CATI surveys provide accurate data at a reasonable cost..

2. **Define the unit of analysis**: The unit of analysis is the level at which final estimates will be made. Some studies have multiple units of analysis: process evaluation results may be based on respondent level estimates, while impact evaluation results may be based on measure or project level estimates. Sampling units do not need to be the same as the unit of analysis, but identifying both early is crucial.

For these programs we are recommending a unit of analysis that is a measure type at a site in a given year. We are using the same definition for our sampling unit. Most customers have no more than three projects in a given year, and most projects are of only one or two measure types, so that we will be able to inquire about all of these in a single survey or interview of reasonable length.

We plan to ask each sampled customer about attribution for all of the customers measures. The handful of customers with more than three measures will be included in the IDI sample frame in order to handle their complex projects.

For customers with large numbers of projects and measures, we will ask about groups of measures or projects. The groupings will depend on details of the types of measures and savings magnitudes. We will select a subsample of three groups with probability proportional to size. Typically this will mean asking about two groups that account for the large majority of savings, and a third smaller one selected at random.

3. **Identify the target population:** The target population is the universe of items that inferences and estimates are desired for. For this evaluation the primary target population is future programs of the same type. Having future program years as the target population has two implications for the sample design. First, the applicable error associated with our estimates is the non-finite population corrected error (described in our discussion of sample size below) which requires larger sample sizes for a given precision. Second, analysis by sub-domains such as measure types within the programs becomes more important. The measure mix in programs changes from year to year and typically NTG varies more across measure types than within. For more accurate estimates of net savings for future program years, applying measure type NTG ratios will be preferred to program as a whole NTG ratios.

4. **Establish the Sample Frame:** The sample frame refers to the list or mechanism from which the sample is drawn. A perfect frame will match the target population exactly.

   Since the target populations of this study are the future programs, we will not have a perfect sample frame; however, if the program designs remain relatively stable, using past program participants as the sample frame will provide a good list from which to draw our sample.

5. **Determine sample size:** Sample size refers to the number of items that are selected from the sample frame in order to draw inferences and create estimates about the target population. In stratified designs, sample sizes are determined for each stratum.

   Critical to the sample size determination is the error ratio for each sampling cell with respect to the ratio to the estimated. The error ratio for ratio estimation is the equivalent of the coefficient of variation for estimation of a population mean. Our experience with conducting similar NTG studies of commercial-industrial customers is that the error ratio for the free rider NTG factor is between 0.7 and 0.8 within reasonably defined sampling cells.

   In determining these sample sizes, the number of customers in the full population is also important for two reasons. First, if we are trying to estimate a parameter for a finite population, the sample size required is reduced by the Finite Population Correction or FPC. Second, we need to consider the number of completed surveys we can realistically complete given likely response rates.

Use of the FPC is appropriate when the parameter of interest represents a particular population. This situation applies when we are determining the free ridership factor or spillover rate for a particular program and time frame. When we determine these factors for all future theoretical projects, it is arguably more appropriate to treat the sample drawn from recent participants as coming from an essentially infinite population. Thus, for projection to future years we generally recommend against applying the FPC.

Because we recommend against applying the FPC and we anticipate response rates of 50 percent for our IDIs and 25 percent for our CATI surveys, we recommend *attempting* a census of participants[14]. This will allow for faster data collection as we will not need to maintain a strict call order and will result in the most completes possible for the data collection methods used.

6. **Stratification:** Stratification is the partitioning of a target population. Stratification is often introduced in a design for two reasons:

   When one desires a specific sample size within small groups of the target population, the groups are often used as a stratification variable. For example, the EAC is interested results by measure type, so measure type is being included as a stratification variable in order to obtain a specific number of selected units in each measure type.

   Stratification is also used in a design to increase precision. A sample design is optimized when strata can be formed so that the variability of the primary and secondary outcome measures are as small as possible within strata and as large as possible between strata. We optimized the sample design's size-based strata (m^3) using a model based optimization algorithm appropriate for ratio estimation. Optimization based on size ensures more data collected from large sampling units, which improves the precision of the estimates.

7. **Sample Selection:** Sample selection refers to the process of obtaining the sample of units from the sample frame. If all units on the sample frame are selected then the design is referred to as a census or certainty sample. Otherwise units may be selected either randomly or nonrandomly, depending on the evaluation goals, constraints and amount of acceptable bias. The sample selection process is a critical feature of the sample design and has a direct impact on the expected precision and bias of estimates. The optimal sample selection process for a particular project can vary greatly.

8. **Unit and Item non-response** Unit and item nonresponse are potential sources of bias, depending on the nonresponse mechanism and the level of nonresponse encountered. Unit nonresponse refers to the absence of information from an entire sampled unit. Item nonresponse refers to the situation where some data are collected, but not all, from a sampled unit. The nonresponse mechanism refers to the process that is causing the nonresponse. If the probability of responding depends on the data items being sought then the nonresponse mechanism is said to be nonignorable. Otherwise it is called an ignorable nonresponse mechanism. Nonresponse bias tends to be greater when the nonresponse mechanism is non-ignorable and as levels of item nonresponse increase.

[14] DNV GL will attempt a census if total size falls within the original scope. If the number of participants provides too great a sample size for the established scope, sampling will be used.

There are various ways to address nonresponse in a sample. For example, weight adjustments are often used to account for unit nonresponse and item imputation techniques are often used to account for item nonresponse.

If nonresponse levels are low and the response mechanism is thought to be ignorable then one could ignore nonresponse and simply create estimates among the respondents.

We recommend treating unit nonresponse as ignorable for this study since it does not depend on the data items being sought. Instead, it depends on the willingness of the decision maker at the participating business agreeing to respond to the survey.

For item nonresponse in the scored portion of the surveys we recommend treating the nonresponse as nonignorable if all three of the T,E,Q portions of the free ridership sequence contain non-response. Otherwise we plan to treat the item nonresponse as ignorable and will impute the average response for the missing item from among scored units of the same measure type and utility. The exception to this rule is when we find conflicting responses in our QC of the data collection that indicates the nonresponse is nonignorable. For nonignorable item nonresponse we will drop the unit from the analysis.

9. **Expansion** Sample expansion refers to the process of extrapolating results from a sample back to the target population of interest. Often times this is done using a sample weight. The weight is a numeric quantity associated with each responding unit and conceptually represents the amount of the target population the responding unit represents during the analysis. The sample weight is some function of the total number of units on the sample frame.

   The sample weight for our analysis will be built from the inverse probability of selection, incorporating additional adjustment factors to account for nonresponse and coverage errors. The sample weight will be utilized along with the "size" of the unit (energy savings) to expand results using ratio estimation, as described in the ratio estimation appendix of this work plan.

10. **Domains of interest:** Often times, estimates for an entire target population are of interest, but so are estimates for various subgroups. Subgroups may or may not overlap. Identifying the population domains of interest is another critically important design feature because it affects the decisions being made about other design features, such as the desired sample size, stratification variables and primary and secondary estimates of interest.

We are providing the EAC with estimated precision for domains of interest in the next section of this work plan. After EAC review of the draft work plan, we can add adjust the definitions of the domains of interest to best reflect the level at which the EAC is likely to want results presented in the final report.

APPENDIX B NET-TO-GROSS METHODOLOGY

The Ontario DSM Guidelines define a free rider as “a program participant who would have installed a measure on his or her own initiative even without the program.” [15]

A great deal of attention has been given to the challenges of “scoring” attribution surveys. In DNV GL’s free ridership approach, we use a critical set of lead-in questions to establish the framing, determine that we have the right respondent, and clarify what measure is being asked about. We then have essentially one question each identifying the effect of the program on the efficiency, the quantity, and the timing of the measure installed. We include some cross-checks; if an inconsistency is identified in these checks we probe to resolve that inconsistency.

DNV GL has developed a streamlined and effective approach to these question sequences. While many of the set-up questions are not used explicitly in calculating measure free ridership, our experience is that dispensing with these framing questions, or attempting to compress the scored questions into a more general subset, results in responses of ambiguous meaning.

Other practitioners prefer to ask the same question multiple ways and then average the corresponding responses. We find this approach typically means blending a looser question with a tighter one, and also increases response burden. We prefer to ask each needed element one way, with appropriate framing and wording to ensure meaningful responses, and to probe as needed.

A frequent challenge in this context is how “don’t know” responses are treated. We typically find that with well framed questions addressed to appropriate respondents, “don’t know” responses are rare, on the order of 5 to 10 percent. As a result, our overall estimates are not highly sensitive to how these cases are treated. If a respondent gives a large number of don’t know responses, we treat that survey as essentially uninformative.

The outcome of our free ridership analysis for a particular respondent and measure is the net program-attributable savings for that measure. This net savings takes into account

- The program may get credit for part but not all of the energy efficiency improvement
- The program may get credit for some but not all of the units installed
- The program may get credit for accelerating the timing of the measure implementation, with or without getting credit for increased efficiency or quantity.
- For a complex project, the program may get credit for some elements of the project, and not for others.

This approach treats free ridership as a fraction of gross savings, for both individual measures and for the program as a whole. DNV GL believes this is a more meaningful approach than attempting to classify each participant, project, or measure as a free rider or not. An excerpt from a prior report detailing the survey approach and associated calculation rules are provided in Appendix C.

Likewise for spillover, after collecting information on what additional measures were implemented based on experience with a program-attributable measure, we determine the savings associated with these measures via engineering analysis.

---

[15] Ontario Energy Board *Demand Side Management Guidelines For Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

Interpreting Customer Responses on Program Effect on Timing.

There are two general ways to treat survey responses indicating that the program caused a measure to be implemented sooner than it otherwise would have. DNV GL has tools and procedures for handling both of these approaches, with a range of specific for either.

Scaling based on reported acceleration

One way to treat the statement that the measure would otherwise have been implemented x months or years later is essentially like a scaling or probability factor. If the measure reportedly would have been implemented within a very short time absent the program, it’s highly likely that it would indeed have been implemented. If the measure reportedly would have been implemented a long time out, it’s less certain that it ever would have been implemented. Accordingly, attribution is scaled down if the reported timing acceleration is very short, but scaled down less for greater acceleration. The simplified version of this approach gives no credit if the measure would have been done within say 1 or 2 years, and full credit thereafter. An approach DNV GL has used for multiple programs is to give full credit if the reported acceleration is greater than 4 years, and scale the credit linearly between 0 months and 4 years.

We will be using 4 years as our standard for this project.

Dual baseline

The second general way to account for acceleration is to take the reported acceleration at face value, and recognize a different baseline condition before and after the acceleration period. For example, if old equipment would otherwise have been replaced 2 years later, the baseline for those first 2 years is the old equipment. For the remainder of the measure life, the baseline is the efficiency that would otherwise have been installed at that point.

Even with the dual baseline approach, we recommend disregarding reported acceleration greater than 4 or 5 years, since customer investment plans are unlikely to be determined that far out. With the dual baseline approach, the attribution factor is the ratio of the total net savings over the life of the measure to the total gross savings over that period.

We will not be using the dual baseline approach on project: TEC agreed on using Y1NS method after the kickoff meeting.

Determining Attribution Parameters for the Program.

Once we have determined the program-attributable savings for the program measures for each surveyed customer, we calculate the in-program attribution rate by sample expansion and ratio estimation. That is, we calculate the ratio estimator

NTG_{FR} = (Net Savings)$_{FR}$/(Gross Savings)$_{FR}$
FR = 1-NTG_{FR}

Where

NTG_{FR} = net-to-gross ratio considering free ridership only (not spillover)
(Net Savings)$_{FR}$ = estimated program non-free rider savings, from expanding the FR survey sample
(Gross Savings)$_{FR}$ = estimated program gross savings, from expanding the FR survey sample.
FR = free ridership as a fraction of program gross savings.

Likewise for spillover (SO):

SO = (Spillover Savings)/(Gross Savings)

(Spillover Savings) = estimated total spillover savings, from expanding the spillover survey sample

$(\text{Gross Savings})_{SP}$ = estimated program gross savings, from expanding the spillover survey sample.

Overall NTG is then calculated as

NTG = (1-FR) x (1 + SO).

APPENDIX C SPILLOVER METHODOLOGY

Spillover "refers to effects of customers that adopt energy efficiency measures because they are influenced by a utility's program-related information and marketing efforts, but do not actually participate in the program."[16] As in many jurisdictions, Ontario's Demand-Side Management Guidelines recognize the importance of spillover in determining program benefits, and also require "comprehensive and convincing empirical evidence" to support any program spillover claim.

Key challenges to providing convincing quantified evidence of spillover for a particular customer include:

- Determining that a particular subsequent action was due to the influence of the program
- Confirming that the action was not taken as part of the original or another program, hence already counted by the program
- Quantifying the savings associated with confirmed spillover actions.

DNV GL's proposed approach provides a high level of rigor to address each of these issues.

- We determine program influence using participant surveys that start with the framing of our free ridership questions, and the identification of the influence of the program on the original measures. This framing helps ensure more meaningful responses to questions of the influence of the original measure on subsequent actions. As for the free rider surveys, obtaining the right respondents is also essential to obtaining meaningful responses to these questions.

- We confirm that the actions tentatively identified as spillover were not already counted by another program by cross-checking tracking data bases. Also critical to separation of spillover from program-claimed savings is understanding what savings if any are claimed by the programs for facilitation support such as opportunity identification, feasibility studies, audits, and related continuous improvement program engagement.

- We quantify the savings for confirmed spillover actions by collecting engineering specifications and calculating associated savings. This approach gives more accurate results than asking customers to estimate the magnitude of spillover savings relative to the original measure.

Thus, our participant spillover methodology addresses the following key issues:

- *Locating the right decision-maker* - Large commercial and industrial companies have multiple decision-makers and it is often difficult to find someone who is familiar with both the tracked program-influenced measure and the spillover measure. Employee turnover can also complicate this. Our approaches to ensure appropriate respondents are discussed above.

- *Avoiding double-counting* – Companies that received financial incentives from an energy efficiency program for one measure are likely to seek these incentives for future measures. Hence it is important to get the program's latest tracking data to make sure that a potential participant spillover measure did not receive program support.

- *Estimating program attribution for potential spillover measures* - A common way of assessing participant spillover is to ask how much the participant's experience with the tracked program-influenced measure influenced their decision to implement measures that are candidates for spillover

[16] Ontario Energy Board *Demand Side Management Guidelines For Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

attribution. It is difficult to turn this "fuzzy" assessment of program influence into a more concrete attribution factor necessary for attributing a certain quantity of therms from the spillover measure to the program.

- *Estimating the energy savings for the participant spillover measures.* Because spillover measures occurred outside the program, evaluators do not have access to the same information about the size, type, and quantity of the implemented energy-efficient measures that they would find in a program tracking database.

Our approach to these issues is described in more detail below. This approach is based on one we used successfully in Wisconsin C&I programs over many years.

Understanding Energy-Related Standard Practices

The first objective of the survey will be to find out whether the participant's company or organization had installed any energy-efficient equipment or made any energy-efficient changes in operation or maintenance (O&M) procedures after the implementation of the tracked project. But before doing that we will collect some information about the company's or organization's energy-related decision-making process. We will ask the participants a series of questions about:

- Who in their company makes decisions about equipment replacement and retrofit projects;
- What information sources are used in making these decisions; and
- Possible barriers to energy efficiency implementation.

This information will be valuable for a number of different reasons. First, it should help program implementers devise strategies for increasing program awareness and mitigating barriers to project implementation, especially for participants who did not identify any subsequent energy-efficient projects after the tracked project. Second, by shedding light on the project decision-making process, it should help the evaluators make better judgments about assigning program attribution to a given project. Finally, it should make the survey appear less peremptory for those who did not report any new energy-efficient projects after the tracked projects, since otherwise their survey would be terminated fairly quickly.

After we collect this information about participant energy practices, we will ask the participants whether their company/organization had installed any energy-efficient projects after the installation of the tracked project. If the participants report no subsequent actions, we will terminate the survey since there is no participant spillover to be measured. If they do identify subsequent projects, then we will collect some basic information about the project including:

- The approximate year of the project;
- The geographic location of the project (e.g. city);
- The types of energy-efficient measures installed or energy-efficient O&M practices implemented; and
- Whether the tracked project and the subsequent project were in the same facility or not (needed for the calculation of inside vs. outside spillover).

Because this information will be collected by CATI program surveyors who do not have an energy background, we will not ask them to try to collect too detailed information about the energy-efficient project. It just needs to be detailed enough to allow the evaluators to make a reasonable match with any projects in the program tracking data.

Calculating Program Attribution for Candidate Spillover Actions

The next stage of the survey will focus on program attribution. Our method only awards spillover energy savings if two criteria are met:

1. The original tracked project is at least partially attributable to the program (Attribution Factor A).
2. The subsequent project is at least partially attributable to the participant's experience with their earlier tracked project (Attribution Factor B).

If a measure met these two criteria, we assign it spillover savings according to the following formula.

(Spillover Savings) = (the measure's annual savings) X (Attribution Factor A) X (Attribution Factor B).

We apply both Attribution Factor A and Attribution Factor B because if the program had no influence on the original tracked project, the program should not get credit for any additional measure installations resulting from that tracked project. To reduce respondent fatigue, if Attribution Factor A is zero (a total free rider) we will not ask them the questions for calculating Attribution Factor B.

To determine attribution factor A we will use the NTG question battery already described in this proposal. For Attribution factor B we will use a scoring method that will be triggered off the question, "If you had not made the earlier energy-efficiency improvements I just listed, how likely would you have been to make this additional energy efficiency improvement?" The scoring method, which we used in Wisconsin for many years, is shown in Table C-1. If the participant said they were very likely to have made the additional energy efficiency improvement without the program, then we will terminate the survey since there will be no participant spillover to be measured.

Table C-1: Program Attribution for Subsequent Measures

If had not made tracked program-influenced energy efficiency improvement, reported likelihood of making subsequent energy efficiency improvement	Assigned Attribution Factor B
Very likely	0.00
Somewhat likely	0.55
Not very likely	0.90
Very unlikely	1.00

The reason we use a different method for Attribution Factor B than for Attribution Factor A is that the character of influence is different. For the program's influence on the tracked project (Attribution Factor A) financial incentives usually account for much of the influence in terms of reducing payback periods and therefore we want to measure things like acceleration effects. However, with participant spillover the influence is less tangible and more likely to be general positive experience with a new energy-efficient technology and the energy savings it produces. We believe that using a Likert scale question will better capture the less tangible character of this type of influence.

Avoiding Double Counting of Energy Savings

Once a participant has identified a subsequent project that is attributable – e.g. one where Attribution Factor A and Attribution Factor B are both greater than zero -- then we will conduct some additional checks to insure that the subsequent project is not also a tracked project. Some of these checks will occur in the survey itself. For example, we will ask the participants if they recall receiving financial incentives from an energy efficiency program for the subsequent projects. We will also examine the program tracking data to make sure that the subsequent project is not in the tracking program data for future years. For example, if we interview a 2011 participant and they identify a subsequent project in 2013 we will look at the 2012-2013 program tracking data (we will look at both program years in case their memory of the project timing was faulty) to see if we can find that project. If we do find the subsequent project in program tracking data, then we will remove that project as a candidate for spillover energy savings since the savings for that project has already been claimed by the program.

Estimating Energy Savings for Participant Spillover Measures

Once a project has been identified as having spillover energy savings (it is program attributable and we could not locate it in the program tracking data) the final step will be to estimate its energy savings. To estimate the annual energy savings for participant spillover measures, we plan to have engineers conduct follow-up interviews with the persons identified in the CATI surveys as being most familiar with the spillover projects. The engineers will have some basic project information collected from the CATI survey as well as some information about deemed savings algorithms for that measure which will allow them to prepare ahead of time the types of questions they will need to ask (e.g., about baseline measures, hours-of-use, etc.). Once they have conducted the interview and collected the necessary information they will calculate the first year savings for the measure. If a deemed savings algorithm exists for that measure they will use that as a default. If none exists then they will use their best professional judgment to estimate the energy savings.

The final step will be separating the spillover energy savings estimates that were reported for “inside” facilities vs. those reported for “outside” facilities. These savings estimates will then be used to produce the inside and outside spillover energy savings estimates for the report.

APPENDIX D PROJECT TASKS AND SUBTASKS

Task 1: Project Kickoff

- ☒ **Task 1.1: Convene a project kickoff meeting**
- ☒ **Task 1.2: Reach a consensus on methodology**

Task 2: Sample Design

- ☐ **Task 2.1: Data Exploration**
  - ☒ Initial data exploration , Union and Enbridge
  - ☐ Exploration of the full datasets, Union and Enbridge
- ☐ **Task 2.2: Define the Unit of Analysis**
  - ☒ Initial unit of analysis definition
  - ☐ Final unit of analysis definition using full datasets
- ☐ **Task 2.3: Stratify the Data**
  - ☒ Initial stratification
  - ☐ Final stratification using full datasets
- ☐ **Task 2.4: Design the Sample**
  - ☒ Initial sample design
  - ☐ Full sample design and precision optimization
- ☐ **Task 2.5: Prepare the Sample Frame**
  - ☐ Sample contact information and documentation request
  - ☐ Prepare the sample frame

Task 3: Data Collection

- ☒ **Task 3.1: Program Managers and Staff Interviews**
  - ☒ Union Gas Program Portfolio Management
  - ☒ Enbridge Commercial Program Managers
  - ☒ Enbridge Industrial Program Managers
  - ☒ Enbridge Run-It-Right Program Managers
- ☐ **Task 3.2: Program Energy Solutions Consultants Interviews**
  - ☐ Energy Solution Consultant Interview Guide
  - ☐ 10 pre-survey interviews
  - ☐ 10 follow up interviews
- ☐ **Task 3.3: Program Participant CATI Survey**
  - ☐ CATI survey Instrument
  - ☐ ## CATI survey interviews completions attempted
- ☐ **Task 3.4: Program Participant In-Depth Interviews**
  - ☐ In-Depth Interview Instrument, mirroring CATI instrument
  - ☐ ## IDI completions attempted
- ☐ **Task 3.5: Program Participant Engineer Interviews**
- ☐ **Task 3.6: Participating Vendor In-Depth Interviews**
  - ☐ In-Depth Interview Instrument
  - ☐ 80 IDI completions attempted

Task 4: Data Analysis Subtasks
☐ **Task 4.1: Analyze survey and interview data**
☐ Sample frame data transformation
☐ Sampling weight
☐ Data validity check
☐ **Task 4.2 Calculate estimates**
☐ Attribution
☐ Spillover
☐ Net-to-Gross
Task 5: Reporting Subtasks
☐ **Task 5.1: Monthly Status Reports**
☐ **Task 5.2: Bi-Monthly Updates**
☒ **Task 5.3: Methodology Memo**
☐ **Task 5.4: Draft report**
☐ Include estimates of free ridership
☐ Include estimates of participant spillover
☐ Include forward free ridership and spillover data
☐ **Task 5.5: Final report and presentation**
☐ Final report addressing comments on draft report
☐ In-person presentation
Task 6: Project Management
Task 6.1: Complete evaluation on time, on budget and within scope.
Task 6.2: Keep client informed

DNV·GL

MEASUREMENT OF NTG FACTORS AND CUSTOM SAVINGS VERIFICATION FOR ONTARIO'S NATURAL GAS CUSTOM COMMERCIAL AND INDUSTRIAL DSM

Scope of Work

for Ontario Energy Board

Date: December 14, 2016

Table of Contents

OVERVIEW

This document has been prepared for the Ontario Energy Board (OEB) and provides the scope of work for the combined Custom Program Savings Verification (CPSV) and Net-to-Gross (NTG) Evaluation of Enbridge Gas Distribution Inc.'s (Enbridge) and Union Gas Limited's (Union) natural gas demand-side management (DSM) programs delivered in 2015. The combined study will produce verified savings, free ridership (FR) and spillover (SO) ratios. Whereas the CPSV and FR results will rely on 2015 program results, the SO results will be based on data collected from 2013 and 2014 program participants. Results from the SO study will be applied to the 2015 program results (along with the CPSV and FR results) in the calculation of verified net savings. Projects included in each portion of the study are shown in Table 1.

Table 1: CPSV, FR, and SO by program

Program		2015		2013/14
		CPSV	FR	SO
Union				
Custom	Large Volume	✓	✓	✓
	Commercial & Industrial*	✓	✓	✓
	Low Income Multi-Residential	✓		
Enbridge				
Custom	Commercial*	✓	✓	✓
	Industrial	✓	✓	✓
	Low Income Multi-Family	✓		
RunitRight			✓	✓

*Custom Market-Rate Multi-Residential projects are included as a part of this program.

Evaluation Objectives

The overall goals of the combined evaluation are to develop transparent

1. verified gross and net savings for 2015 custom commercial, industrial and large volume projects
2. free ridership rate for Enbridge's 2015 RunitRight program
3. spillover factors applicable to custom commercial, industrial and large volume projects and RunitRight, based on projects claimed in 2013 and 2014 and applicable to projects installed in 2015 and future program years

Evaluation Approach

The methodology selected for the CPSV portion of the study consists of engineer reviews of gross savings. Reviews of complex projects will include on-site verification and data collection, while less complex projects will be verified with Telephone Supported Engineering Reviews (TSERs).

The methodology selected for the NTG evaluation will rely on end-user self-report surveys and interviews. The end user self-reports will be supplemented by project-specific interviews with vendors to capture indirect effects of the program on end-user decision making. Surveys and interviews will be collected from the most recent program years in order to create NTG factors that will be most meaningful for future years.

Deliverables and Schedule

This study will result in three final deliverables:

1. 2015 CPSV and Free ridership Report
2. Spillover Report
3. Final 2015 verified and net savings memo

The current project schedule is shown in Table 2. Because data collection is schedule to fall during the holiday season, we included four weeks of additional time for those tasks.

Table 2: High Level Schedule

Task	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
SOW										
Data Collection Planning										
Data Collection										
Analysis										
Reporting						D (1)	F (1)		D (2,3)	F (2,3)
Project Management										

D. Draft
F. Final

1. 2015 CPSV and Free ridership Report
2. Spillover Report
3. Final 2015 verified and net savings memo

Key Concepts and Terms

This section defines several key concepts that will be used throughout this work plan, using the definitions from the Ontario DSM Guidelines for spillover and free rider.

- *Spillover* "refers to effects of customers that adopt energy efficiency measures because they are influenced by a utility's program-related information and marketing efforts, but do not actually participate in the program." [1] We consider both inside and outside, and both like and unlike spillover through this project.
  - *Inside spillover* "refers to non-incented measures that were installed within the same project or facility."[2]
  - *Outside spillover* "refers to measures for which the customer did not receive an incentive adopted in an outside location or unrelated project for a participating customer." [3]
  - *Like spillover* refers to non-incented measures of the same type as incented measures.[4]
  - *Unlike spillover* refers to non-incented measures of a different type as incented measures[5]
- A *free rider* is "a program participant who would have installed a measure on his or her own initiative even without the program."[6]
- *Gross savings* are "the changes in energy consumption and/or demand that result directly from program-related actions taken by participants in an efficiency program, regardless of why they participated."[7]
- *Net savings* are "the changes in energy consumption or demand that are attributable to an energy efficiency program. The primary, but not exclusive, considerations that account for the difference between net and gross savings are free riders (i.e., those who would have implemented the same or similar efficiency projects, to one degree or another, without the program now or in the near future) and participant and non-participant spillover[8] (i.e., savings that result from actions taken as a result of a program's influence but which are not directly subsidized or required by the program). Net savings may also include consideration of market effects (changes in the structure of a market)."[9]
- The *net-to-gross* (NTG) ratio is an adjustment factor that reduces savings due to free ridership and increases savings to account for spillover. The NTG ratio "is the portion (it can be less than or greater than 1.0) of gross savings (those that occur irrespective of whether they are caused by the program or not) that are attributed to the program being evaluated."[10] The NTG ratio is a combination of NTG factors that include the spillover and free ridership rates.
- *Attribution* is the portion of a measure that is attributable to the program. For program measures attribution is the complement of free ridership (1-FR).
- *Energy Advisors* are utility and/or program staff who provide information to customers about energy saving opportunities and program participation. This is a general term that includes, but is not limited to Enbridge's Energy Solutions Consultants and Union's Account Managers.

---

[1] Ontario Energy Board *Demand Side Management Guidelines for Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

[2] Ontario Natural Gas Technical Evaluation Committee (TEC), *Request for Proposal: Measurement of Net-to-Gross (NTG) Factors for Ontario's Natural Gas Custom Commercial and Industrial Demand Side Management (DSM) Programs,* RFP-002-2013 (2), December 2013, Section 2.

[3] Ontario Natural Gas Technical Evaluation Committee (TEC), *Request for Proposal: Measurement of Net-to-Gross (NTG) Factors for Ontario's Natural Gas Custom Commercial and Industrial Demand Side Management (DSM) Programs,* RFP-002-2013 (2), December 2013, Section 2.

[4] NREL, Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures, December 2014. http://www.nrel.gov/docs/fy14osti/62678.pdf

[5] Ibid

[6] Ontario Energy Board *Demand Side Management Guidelines for Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

[7] SEE Action, Energy Efficiency Program Impact Evaluation Guide: Evaluation, Measurement, and Verification Working Group, DOE/EE-0829, December 2012. https://www4.eere.energy.gov/seeaction/sites/default/files/pdfs/emv_ee_program_impact_guide_1.pdf, page xiv

[8] Note: Non-participant spillover is not within the scope of this study.

[9] Ibid

[10] Ibid, page 5-1

- *Vendors* are program trade allies, business partners, contractors and suppliers who work with program participants to implement energy saving measures.
- *Computer-Aided Technical Interviews* (CATI surveys) are structured surveys administered by a third-party survey firm. CATI surveys require clearly defined skip logic and structured formats. CATI surveys are a lower cost data collection approach that is suitable for structured gathering of information from large samples of respondents.
- *In-Depth Interviews* (IDIs) are structured technical interviews administered by evaluation engineers and market researchers either in person or more frequently, over the phone. IDIs offer more flexibility than CATIs and are best leveraged for complex projects and topics.

Description of Included Programs

The programs included in the evaluation include portions of the utilities' resource acquisition, low income, market transformation, performance-based and large volume portfolios.

Low Income Multi Residential Affordable Housing (Enbridge)
Low-Income Multi-Family Offering (Union)

The programs offer multi-family low income housing customers with incentives to encourage energy efficient upgrades and funding for energy audits. The programs also provide technical services, benchmarking, and education for housing providers, building operators and tenants about their building's energy usage and ways to achieve energy efficiency. Eligible measures include boilers, ventilation systems, building envelope, window upgrades, in-suite water conservation measures (faucet aerators, showerheads), and heat reflector panels.

The target markets for both programs are social and assisted housing providers who own and operate Part 3[11] buildings and private multi-residential building owners that provide housing to low income households. In addition, Enbridge targets shelters and supportive housing.

In this Scope of Work we refer to these programs collectively as Low Income Multi-Family (LI MF).

Custom projects implemented as part of these programs and claimed in 2015 are included in the CPSV portion of the study.

The NTG (FR and SO) evaluation portion will not look at projects implemented as part of these programs.

The program metrics for the full program and the cumulative savings included in the 2015 evaluation are shown in Table 3. Ten percent of the combined LI MF program savings are from custom projects, 4% of Enbridge's program and 33% of Union's program.

Table 3: Low Income Multi-Family 2015 Scorecard Metrics

Low Income Multi-Family	Budget	Savings	S. incentive	Budget	Savings	S. incentive	Budget	Savings	S. incentive
	Enbridge			Union			Combined		
Scorecard Metrics	$ 2,111,746	69,226,782	$1,223,674	$ 2,271,917	16,965,778	$1,316,926	$ 4,383,663	86,192,560	$2,540,600
Custom Projects included in 2015 CPSV and FR Evaluation		63,801,575			5,624,627			69,426,202	

[11] "Part 3" references buildings covered by Part 3 of the Ontario Building Code, defined as those exceeding 600 square meters in area or greater than three storeys in height; for residential energy efficiency programs, these are typically multifamily buildings.

Large Volume (Union)

Union continues to encourage the adoption of energy efficient equipment, technologies, and actions through direct customer interaction via its Large Volume program. The Large Volume program is applicable to customers in Rate 1 (2015 only) and Rate T2/Rate 100.

The 2015 to 2020 program uses a direct access budget mechanism for the customer incentive budget process for Rate T2/Rate 100 customers. This mechanism grants each customer direct access to the customer incentive budget they pay in rates. Customers must use these funds to identify and implement energy efficiency projects, or lose the funds which will consequently become available for use by other customers in the same rate class. This “use it or lose it” approach ensures each customer has first access to the amount of incentive budget funded by their rates. The incentive approach for Rate T1 customers remains unchanged from the aggregate pool approach offered in 2014.

The Large Volume program is the only “direct access” program offered in Ontario. It is similar in concept, though not in funding mechanism design, to the standard custom programs offered by the two gas utilities and to the electricity CDM Process and Systems program offered by electricity distributors. It also overlaps to some extent with the Custom Track of the electricity CDM Retrofit program.

Custom projects implemented as part of this program and claimed in 2015 are included in the CPSV and FR portions of the study.

Custom projects implemented as part of this program and claimed in 2013 or 2014 are included in the SO portion of the study.

The program metrics for the full program and the cumulative savings included in the 2015 evaluation are shown in Table 4.[12] The table shows that while most of the Large Volume is Custom and falls within the scope of this evaluation, a small percent of savings (<1%) come from prescriptive projects.[13]

Table 4: Large Volume 2015 Scorecard Metrics

Large Volume	Budget	Savings	S. incentive	Budget	Savings	S. incentive	Budget	Savings	S. incentive
	Enbridge			Union			Combined		
Scorecard Metrics				$ 3,209,716	578,023,195	$ -	$ 3,209,716	578,023,195	$ -
Custom Projects included in 2015 CPSV and FR Evaluation					575,404,661			575,404,661	

Commercial and Industrial Custom Program (Enbridge & Union)

The custom program offerings have been designed to encourage commercial and industrial customers to reduce their energy consumption by providing customer-specific energy efficiency and conservation solutions. The custom programs provide financial incentives, technical expertise, and guidance with respect to energy related decision making and business justification, including helping customers to prioritize energy efficiency projects against their own internal competing factors and demonstrate the competitive advantage customers can gain through efficiency upgrades. These custom programs differ from the prescriptive and direct install programs as they provide tailored services and varying financial incentives based on overall natural gas savings realized by the customer to address customer-specific needs. The custom programs build upon

12 Cumulative savings included in the evaluation are based on project data sent by Union on August 8, 2016.

13 Union Gas provided the savings from and counts of prescriptive projects that were claimed as part of the Large Volume program via email May 31, 2016.

those deployed by the gas utilities in past. They are very similar to, and serve effectively the same customers as, the electricity CDM Retrofit Program's Custom Track.

The goal of the Enbridge Commercial Custom offer is to reduce natural gas use through the capture of energy efficiency opportunities in commercial buildings, including retrofits of building components and upgrades at the time of replacement. The offer aims to promote the highest level of energy efficiency.

The Enbridge Industrial Custom Solutions offer is designed to capture cost-effective energy savings within the industrial sector by delivering customized energy solutions aimed at supporting customers through a continuous improvement approach. Industrial Energy Solutions Consultants (ESCs) focus on assisting customers with the adoption of energy efficient technologies by overcoming financial, knowledge or technical barriers.

Union focuses on advancing customer energy efficiency and productivity by providing a mix of custom incentives, education and awareness to C&I customers across all segments. The objective of the Custom offering is to generate long-term and cost effective energy savings for Union's customers.

The Union Custom offering covers opportunities where energy savings are linked to unique building specifications, design concepts, processes and new technologies that are outside the scope of prescriptive and quasi-prescriptive measures. The offering and incentives are targeted directly to the end user, while trade allies involved in the design, engineering and consulting communities assist to expand the message of energy efficiency.

A subset of the projects in these programs is part of the multi-family or multi-residential segment. In this scope of work we refer to these projects as Market-Rate Multi-family (MR MF) in order to distinguish them from the low income multi-family projects.

All projects implemented as part of these programs and claimed in 2015 are included in the CPSV and FR portions of the study.

All projects implemented as part of these programs and claimed in 2013 and 2014 are included in the SO portion of the study.

Table 5 shows the 2015 scorecard metrics and the cumulative savings included in the CPSV and FR evaluation of 2015.[14]

Table 5: Custom C&I 2015 Scorecard Metrics

C&I Custom	Budget	Savings	S. incentive	Budget	Savings	S. incentive	Budget	Savings	S. incentive
	Enbridge			Union			Combined		
Scorecard Metrics	$ 5,489,284	556,659,946	$4,322,644	$ 7,297,352	678,002,610	$3,348,014	$12,786,636	1,234,662,556	$7,670,657
Custom Projects included in 2015 CPSV and FR Evaluation		556,241,778			678,002,610			1,234,244,388	

Run it Right (Enbridge)

Both Enbridge and Union offer similar building optimization programs that are focused on improving operational efficiency among commercial customers.

[14] Cumulative savings included in the evaluation are based on project data sent by Enbridge on August 4, 2016 and Union on August 8, 2016. Enbridge's updated data removed two projects which account for the difference in savings shown.

Through its program, RunitRight, Enbridge provides customers with an energy assessment, technical and implementation assistance and performance monitoring, while Union offers customers incentives for undertaking low or no-cost energy improvements and activities in their facilities through its RunSmart Program.

The SO portion of the study will include projects implemented as part of the RunitRight program in 2013 and claimed in 2014. The FR portion will evaluate projects implemented in 2014 and claimed in 2015.

RunitRight is not part of the CPSV scope for the verification of 2015 projects and is the only program with non-custom projects included in the scope of the evaluation.

RunSmart is not included in this study.

Table 6 shows the 2015 scorecard metrics and the cumulative savings included in the FR evaluation of 2015.[15]

Table 6: RunitRight 2015 Scorecard Metrics

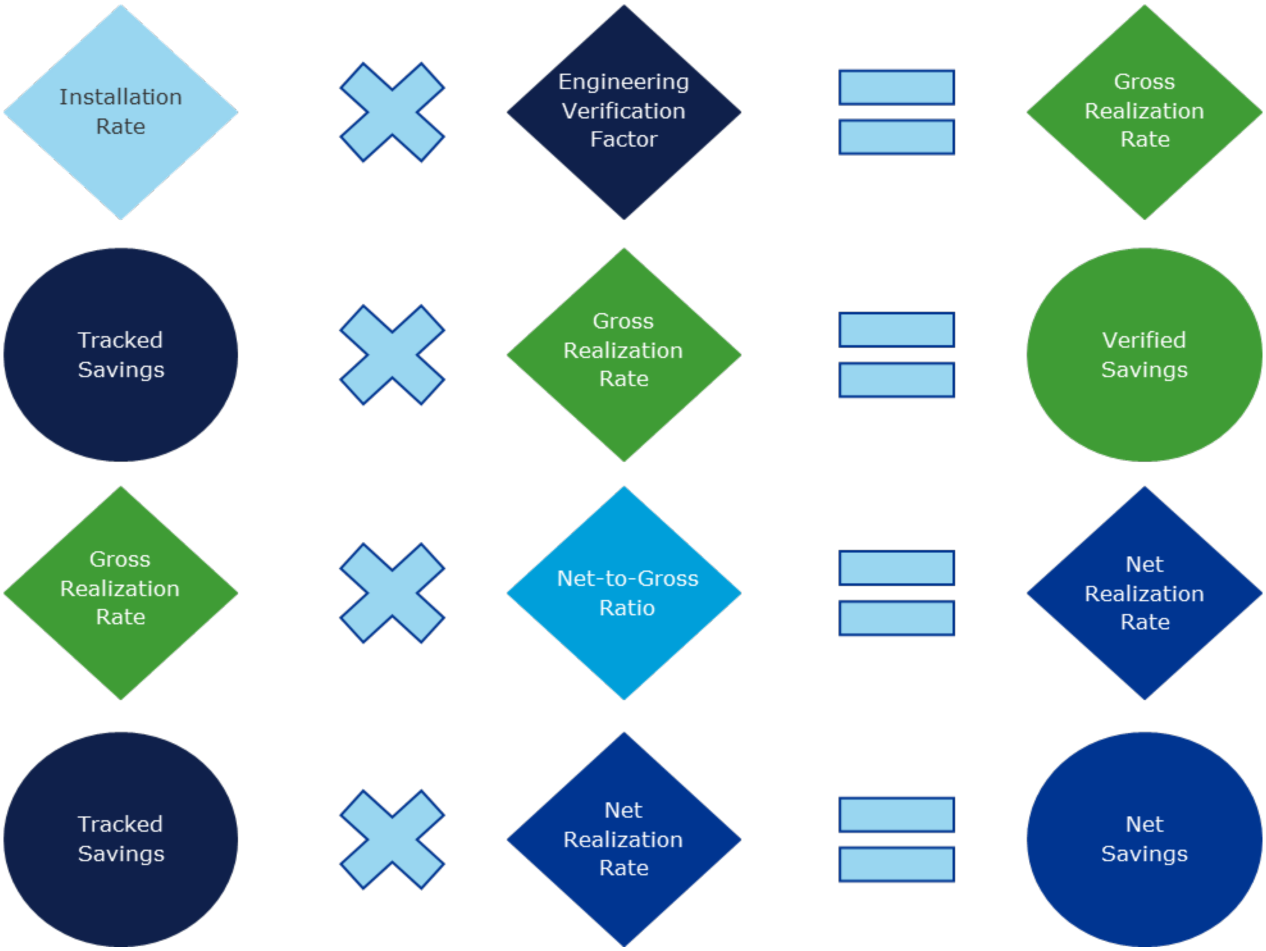
	Budget	Savings	S. incentive	Budget	Savings	S. incentive	Budget	Savings	S. incentive
Run-it-Right	Enbridge			Union			Combined		
Scorecard Metrics	$ 1,181,403	2,684,105	$ 20,843				$ 1,181,403	2,684,105	$ 20,843
Projects included in 2015 FR Evaluation		2,684,105						2,684,105	

Methodology

The overall methodology combines the efforts of the CPSV and the NTG analysis into a single adjustment factor, called the net savings realization rate (Net RR), that can be applied to the reported savings data (or tracked savings) to produce the verified net savings. Figure 1 shows the process for calculating the net RR from the gross savings realization rate (Gross RR) and the NTG ratio, and how it is applied to the tracked savings to produce net savings. The figure also shows the development of the gross RR from the installation rate and engineering adjustment factor, and how it is applied to the tracking savings to produce the verified gross savings.

[15] Cumulative savings in the evaluation are based on project data sent by Enbridge on August 4, 2016.

Figure 1: Ratios used to estimate verified and net savings



At its heart, the analysis is built on three unique adjustment factors, which ultimately combine to produce the gross RR and net RR. The three unique factors are:

- **Installation rate.** This factor corresponds to the fraction of measures that were installed. Each measure is assigned a binary factor that identifies whether it was installed or not installed. Adjustments to the number of units installed for a particular measure are included in the engineering verification factor, not in the installation rate.
- **Engineering verification factor.** This is the ratio of the verified gross savings to the tracking estimate of gross savings for installed measures. The engineering verification factor includes corrections to the numbers of units installed, changes in operating hours, changes in operating levels, etc.
- **Attribution factors.** These factors (which include FR and spillover) are used to determine the proportion of the verified gross savings attributable to the program. The attribution factors are determined from the participant's responses to a battery of survey questions designed to determine how influential the program was in the decision to install a particular measure.

The next sections describe the process used to develop the gross RR (from the installation rate and engineering adjustment factor) and the NTG ratio (from the attribution factors) in greater detail. They also describe the process for expanding the results of the sample to the population, and the methodology for estimating spillover savings and adjustment factors.

Gross Realization Rate

The gross RR is developed through data collected during the CPSV effort, which will verify program-achieved gross savings for measures at a sample of sites. The two components are the installation rate and the engineering verification factor.

- The installation rate is derived through the participant survey data collection, which confirms that the reported equipment / measure or something like it was installed at the facility. The resulting analysis value is binary; any similar project to the one reported is considered installed. At the individual measure level, the installation rate is either 100% or 0%.
- The engineering verification factor is derived from the data collected during the participant survey data collection for TSER projects and through the onsite visits for other projects. Differences between the reported measure and the "substantially similar" measure installed at the facility are accounted for here. The engineering adjustment factor is the ratio of the evaluator-verified savings to the program-reported savings.

The majority of the CPSV process involves determining the evaluator-verified savings estimate for each measure. The measure-level results are then combined using weights from the sample design to an overall adjustment factor.

To get the evaluation-verified savings for each evaluated measure, the CPSV effort will verify savings based on the applicable standard program baseline and measure life based on the best available information. The formula for estimating measure level verified savings is shown here:

$$VGS_L = VY_L \times VGS_S$$

Where:

VGS_L – Verified Gross Savings versus standard efficiency equipment on the market (lifetime)

VY_L – Verified Estimated Useful Life of the equipment/action

VGS_S – Verified Gross Savings versus standard efficiency equipment on the market (annual)

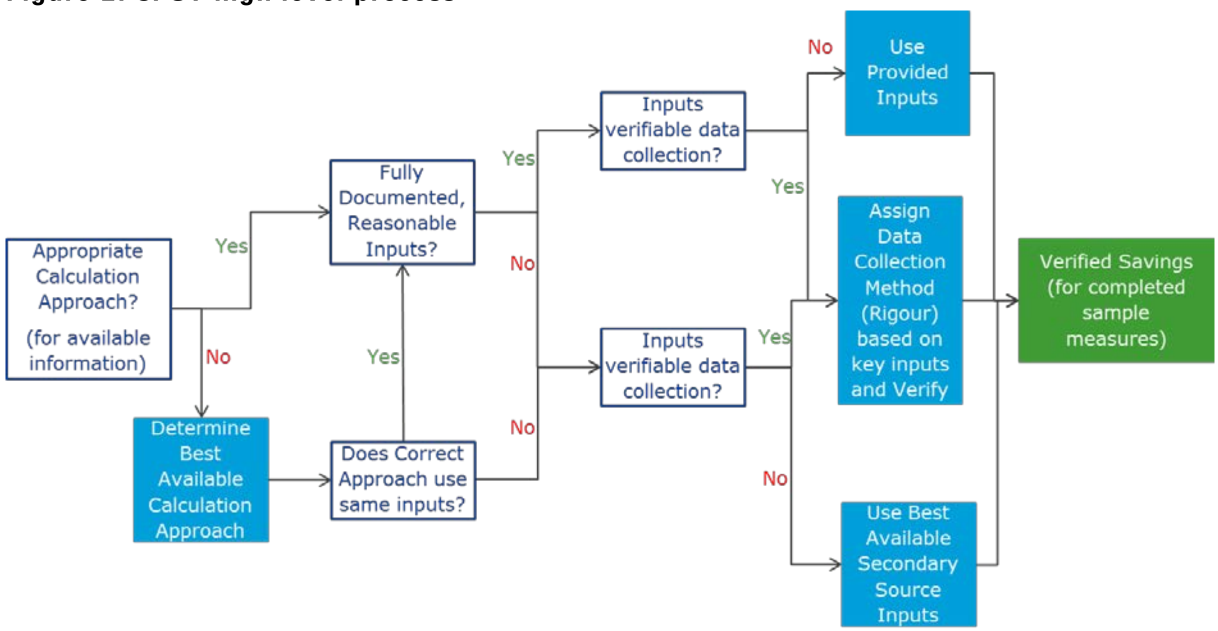
In the Life-Cycle Net Savings (LCNS) method used for this evaluation, the CPSV will also produce a verified savings estimate for accelerated measures using the pre-existing equipment as the baseline (VGS_E). Whether or not the measure is accelerated depends on the responses to the attribution survey and will be discussed later. The "versus existing" verified savings will be used in estimating net savings and will not be included in the verified gross savings. The LCNS methodology is further explained in Appendix B.

The CPSV will produce verified values for three required inputs in the Life-Cycle Net Savings (LCNS) attribution:

- **VGS_S** – Verified Gross Savings versus standard efficiency equipment on the market
- **VY_L** – Verified Estimated Useful Life of the equipment/action
- **VGS_E** – Verified Gross Savings versus existing equipment configuration at the time of installation/action: for a sub-set of measures that are accelerated

CPSV site reports will be completed by assigned evaluation engineers and reviewed by an experienced evaluation engineer at another partner firm. Each review will follow the same basic process shown in Figure 2.

Figure 2: CPSV high level process



After the initial review and savings calculation, an engineer from a partner firm on the EC team (either DNV GL or Itron) will review the site report, approach, calculation, and verified savings. Following this review the verified savings, verified estimated useful life, reasons for deviation and other pertinent information will be compiled into a single dataset at the unit of analysis level for expansion and integration with the FR analysis.

NTG Ratio

The NTG ratio is developed primarily through the data collected from participant and vendor interviews. Data from the engineering verification will also inform the NTG ratio (for the lifecycle net savings (LCNS) approach to free ridership). Where possible, all FR data will be collected via IDIs prior to onsite visits.

The two components of the NTG Ratio are the free ridership and the spillover rates.

- Free ridership (FR) represents the program's influence on the participant's decision to install the measure that received an incentive through the program.
- Spillover represents the program-influenced measures that were installed at the facility as a result of their experience with the original measure. Spillover measures do not receive an incentive.

FR is made up of three factors related to efficiency, quantity and timing. All three attribution factors are based on responses to the attribution questions in the impact evaluation survey. The following is a brief description of each factor:

- **Efficiency attribution, A_E**, measures the effect the program had on the *efficiency* of the equipment installed. The efficiency attribution measures the proportion of savings attributable to the program for increasing the efficiency of the equipment above what would have been installed otherwise.
- **Quantity attribution, A_Q**, measures the effect the program had on the *size or amount* of the equipment installed. The quantity attribution measures the proportion of savings attributable to the program for increasing or decreasing the quantity of equipment above or below what would have been installed otherwise.

- **Timing attribution, A_T**, measures the effect the program had on *when* the equipment was installed. In the LCNS approach the timing attribution is a function of:
  - **Acceleration Period, Y_a**, which corresponds to the number of years between when the equipment was actually installed and when it would have been installed in the absence of the program
  - **Acceleration Period Gas Savings (VGS_E)**, which are estimated versus the pre-existing equipment configuration rather than versus standard efficiency on the market or code. For CPSV sample, this component is calculated as part of the CPSV process and will be included EC team reviews for quality control. This component will not be included in review steps that include OEB or EAC reviewers for respondent confidentiality reasons. For non-CPSV sampled projects in the FR sample, this component will be estimated through using an average ratio from the CPSV sampled projects or based on the age of the existing equipment

The measure-level survey responses are analyzed using a custom software program that objectively determines the FR components and overall rate (see Appendix B and Appendix C for details on the scoring algorithms used) The program includes quality control checks at multiple points in the process. DNV GL has also established a number of metrics that allow us to identify "questionable" results for further investigation and possible correction (details provided in Appendix C). The output of the software program is the source data for the expansion process.

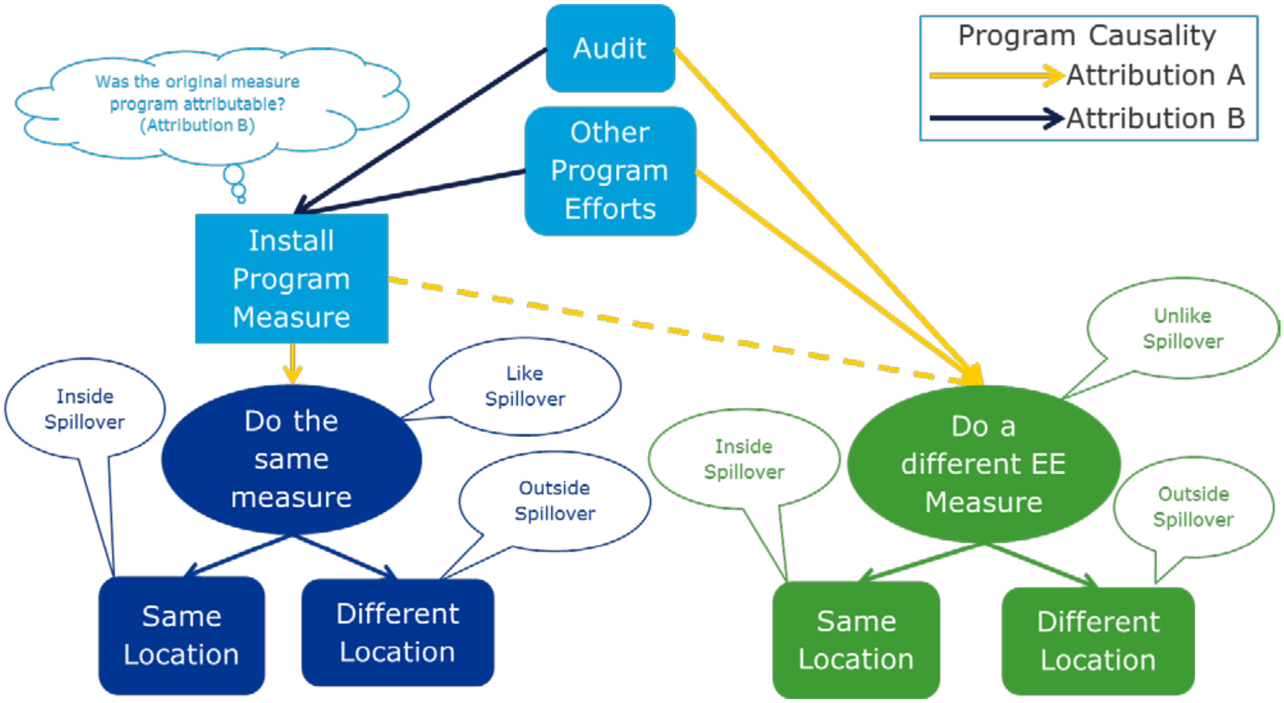
Spillover Estimation

The spillover estimate is developed through data collected from participant and vendor surveys, and a follow-up participant interview. Spillover is present when (see Figure 12):

- A measure is installed after initial program participation
- The measure did not receive an incentive
- The measure was at least partially attributable to the participant's experience with the program in implementing the original measure (Attribution A), **and** (for all like SO and some unlike SO) the original measure is at least partially attributable to the program (Attribution B)

Figure 3 shows how program causality ties to different types of spillover. Attribution B applies to like spillover in all cases, while for unlike spillover, Attribution B applies to the spillover if the original program measure was part of the program influence that led to the spillover measure being implemented.

Figure 3: Program influence on Spillover by Type



Potential spillover projects are identified during an initial survey with the customer and the surveyor collects initial general information on what was installed and whether the new measure was at least partially attributable to the earlier program (attribution A). The analysis team then confirms attribution and compares the participant description with the tracking data for that customer to ensure that the measure did not receive an incentive. If a potential spillover project is identified, a DNV GL engineer will conduct a follow-up phone call to gather the information necessary to estimate the energy savings resulting from that measure, which produces a more accurate savings estimate than asking the customer to provide an estimate themselves. The engineer also will collect the information required for calculating attribution B where it applies.

The relevant attribution estimates are determined based on the information collected during the survey battery and calculated using a custom software program written by DNV GL. The program includes quality control checks at multiple points in the process. DNV GL has also established a number of metrics that allow us to identify "questionable" results for further investigation and possible correction. The program produces measure-level ratios of spillover CCM to tracked or verified CCM, which is the source data for the expansion process.

Sample Expansion

Samples are a necessary part of program evaluation. Sampling reduces costs and customer burden. Nonresponse, whether due to a lack of desire to respond, or because the person that should respond cannot, means that evaluating the entire population usually cannot be done. Any time we evaluate a sample of savings from a program we must expand the sample results to the population. Expanding the results to the population produces results that are representative of the population rather than the sample. Expansion is a key part of calculating important program metrics such as total verified gross and net savings. More detail on sample expansion is provided in Appendix E .

Expansion is done using weights that are determined based on the sample design. The weight is a numeric quantity associated with each responding unit and conceptually represents the amount of the target population the responding unit represents during the analysis. The sample weight is some function of the total number of units in the sample frame. In the CPSV and NTG studies, the sample weight will be built from the inverse probability of selection, incorporating additional adjustment factors to account for nonresponse and coverage errors.

Notation:

N_x = number of units of analysis in stratum X

n_x = number of completed sample units of analysis in stratum X

The weight W_x is calculated as

$$W_x = N_x / n_x$$

The method used to develop the verified or net savings will not affect the weight. In the CPSV, each level of rigour is measuring the same thing (verified savings), only varying in their level of detail. For the NTG portion of the study, the IDI vs CATI distinction operates the same way. In both cases we are looking at energy savings with reliable, valid methods that avoid systematic bias, but with additional magnification on the largest, most variable projects. It is similar to measuring a length using millimetres or eighths of an inch. Both provide accurate measurements of length, but the millimetre measurement is more precise. In terms of expansion, both measurements would get equal weights (once put into comparable units, of course).

DNV GL uses the ratio estimation method to expand our results to the population. The energy saving estimates (tracking savings, installed savings, verified savings or net savings) of the sampled units (measures, projects, sites) are present in both the numerator and the denominator of the ratios, when combined with the sample weights the ratio estimation method produces unbiased, savings weighted adjustment factors. The mathematics of ratio estimation and an example calculation can be found in an appendix.

SUMMARY OF TASKS

The DNV GL team has divided the project into six discrete tasks which are presented, along with their status (as indicated by the box preceding each activity), in Table 7. These tasks are discussed in greater detail in the next sections of the report.

Table 7: Key Project Tasks and Completion Status

Key Tasks
☒ **Task 1: Project Kickoff**
☒ Convene a project kickoff meeting
☒ Reach a consensus on methodology
☐ **Task 2: Sample Design**
☒ Explore the tracking data
☒ Define the unit of analysis
☒ Stratify the data
☒ Design the sample
☒ Select the sample
☐ **Task 3: Data Collection**
☒ Interview program managers and staff
☐ Interview program Energy Advisors
☐ Survey a sample of program participants
☐ Interview a sample of program participants
☐ Onsite verification of a sample projects
☐ Telephone Supported Engineering Review of a sample of projects
☐ Interview program trade allies
☐ Conduct follow-up interviews with program participants
☐ **Task 4: Data Analysis**
☐ Analyze survey and interview data
☐ Calculate estimates
☐ **Task 5: Reporting**
☐ Monthly status reports
☐ Bi-monthly updates
☐ Draft deliverables
☐ Final report and presentation
☐ **Task 6: Project Management**
☐ Complete evaluation on time, on budget and within scope
☐ Keep client informed on progress

We have completed the project kickoff meeting, program manager and staff interviews, and sample design as part of the planning phase, which have informed the specific plan outlined in this document. Next, we will interview utility energy advisors about provide the data collection instruments and associated methodology and will request the contact information and necessary documentation to proceed to the participant data collection phase. The contact request will also ask for vendors associated with the each sampled project. We will calculate the verified savings, free ridership, spillover, and net savings estimates for each program and for domains within programs where there is sufficient sample to provide estimates while protecting respondent confidentiality. These estimates will be provided in the final evaluation reports.

Task 1: Project Kickoff

The initial project kickoff meeting was held on March 17, 2014. At the time, the study did not include CPSV and the client was the Technical Evaluation Committee (TEC). After a long delay, the project was resumed on March 10, 2016 with a reset meeting. Following the reset, oversight of the NTG study was moved from the TEC to the OEB, advised by the EAC. In addition, the January, 2016 OEB DSM Decision included new guidance on how inputs and assumptions (including NTG) for custom programs should be handled in evaluating net impacts.[16] The CPSV scope was added to the project in the months that followed.

Table 8: Task 1 Tasks and Completion Status

Key Tasks	
☒	**Task 1: Project Kickoff**
☒	**Task 1.1: Convene a project kickoff meeting**
☒	**Task 1.2: Reach a consensus on methodology**
☐	Task 2: Sample Design
☐	Task 3: Data Collection
☐	Task 4: Data Analysis
☐	Task 5: Reporting
☐	Task 6: Project Management

[16] Ontario Energy Board, "Decision and Order EB-2015-0029/EB-2015-0049 Union Gas Limited and Enbridge Gas Distribution Inc. Applications for approval of 2015-2020 demand side management plans." January 20, 2016. Page 75.

Task 2: Sample Design

This section presents the stratification plan using the data provided by Union and Enbridge for 2013-2015 custom C&I and multi-family, 2014-2015 RunitRight[17] and 2015 custom Low Income Multi-family projects. Table 9 presents the sample design tasks and their completion status.

Table 9: Task 2 Subtasks and Completion Status

Key Tasks
☒ **Task 1: Project Kickoff**
☐ **Task 2: Sample Design**
☒ **Task 2.1: Explore the Tracking Data**
☒ Initial data exploration, Union and Enbridge
☒ Exploration of the full datasets, Union and Enbridge
☒ **Task 2.2: Define the Unit of Analysis**
☒ Initial unit of analysis definition
☒ Final unit of analysis definition using full datasets
☒ **Task 2.3: Stratify the Data**
☒ Initial stratification
☒ Final stratification using full datasets
☒ **Task 2.4: Design the Sample**
☒ Initial sample design
☒ Full sample design and precision optimization
☐ **Task 2.5: Prepare the Sample and Backup Sample**
☐ Sample contact information and documentation request
☐ Prepare the sample and backup sample
☐ **Task 3: Data Collection**
☐ **Task 4: Data Analysis**
☐ **Task 5: Reporting**
☐ **Task 6: Project Management**

The objectives of this task are to:

- Design representative samples for participant data collection for gross savings verification (CPSV), free ridership (FR), and spillover (SO)
- Achieve 90/10 precision[18] at the desired stratification segment levels (see Table 10):
  - Union FR: three program segments (Custom Large Volume, Custom Commercial, Custom Industrial)
  - Enbridge FR: three program segments (Custom Commercial, Custom Industrial, Run it Right)
  - Union SO: three program segments (Custom Large Volume, Custom Commercial, Custom Industrial)
  - Enbridge SO: three program segments (Custom Commercial, Custom Industrial, Run it Right)
  - Union CPSV: two program segments (Custom Large Volume; Custom Commercial, Industrial, & Multi-family (including market rate and low income))
  - Enbridge CPSV: two program segments (Custom Industrial; Custom Commercial & Multi-Family (including market rate and low income))

---

[17] RunitRight projects claimed in 2014-2015 filings, implemented in 2013-2014.

[18] 90/10 precision refers to 10% relative precision with 90% confidence.

Table 10: Enbridge and Union minimum precision targets by stratification segment

Enbridge			
Stratification Segment	FR Sample Design Rel. Prec.	SO Sample Design Rel. Prec.	CPSV Sample Design Rel. Prec.
Industrial	10%	10%	10%
Commercial	10%	10%	10%
Market Rate Multi-Family			
Low Income Multi-Family	N/A	N/A	
RunitRight	10%	10%	N/A

Union			
Stratification Segment	FR Sample Design Rel. Prec.	SO Sample Design Rel. Prec.	CPSV Sample Design Rel. Prec.
Industrial	10%	10%	10%
Commercial	10%	10%	
Market Rate Multi-Family			
Low Income Multi-Family	N/A	N/A	
Large Volume	10%	10%	10%

This task began with the electronic tracking data and paper documentation submitted by the utilities. The outcome is ordered, stratified samples and backup samples for surveying participants to learn about FR, SO, and verification.

Activities

The actual sample design activities are 2, 3, and 4 in the list below. They are flanked by an exploration of the data to characterize what is available and the actual sample preparation. Each of the activities are described briefly subsequent sections and in detail in Appendix A .

1. Explore the tracking data. This activity ensures that the records provided by the utilities match the records used to develop the reported savings. The activity also characterizes the data in terms of the size of measures, types of measures, and quantity of projects.

2. Define the unit of analysis. The unit of analysis defines the level at which data will be analyzed, but not the level at which it will be collected, which is the sampling unit. It is an important distinction which affects the way the surveys are written, the data is collected, and the domains that can be analyzed from the collected data.

3. Stratify the data. In sample designs, more strata allow the design to control representativeness and estimated precision along more dimensions. Having more strata does not hurt overall precision, but it can increase the sample sizes required. Each stratification level serves to improve efficiency, improve representativeness, or both.

4. Design the sample. In this step, the appropriate coefficient of variation is selected, and the number of targeted data points is determined for each stratum.

5. Prepare the sample and backup sample. The data is organized according to the sample design and an appropriate number of participants are selected as potential study participants.

Deliverables

- Documentation Request
  - CPSV/FR Samples and backup
  - Spillover Sample and backup
- Contact Information Request
  - CPSV/FR Sample and backup
  - Spillover Sample and backup
- Sample Design Appendix for each Report

Task 2.1: Explore the Tracking Data

We explored the tracking data provided by Union and Enbridge to determine data availability, the number and types of measures installed, and the size and quantity of projects. We explored the Union and Enbridge datasets separately. Additional detail can be found in Appendix A .

Enbridge Custom Participant Data

The custom program participant data files provided by Enbridge included custom C&I energy efficiency projects claimed during the 2013-2015 program years and custom Low Income Multi-Family projects claimed in 2015 (Table 11). The records in the tracking data are per installed measure, so there may be multiple rows per project if more than one measure is installed, and there may also be more than one project per account. There are 124 accounts that appear in both the 2015 program year and the 2013/14 program years.

Table 11: Enbridge Custom C&I and Multi-Residential Program Participation Metrics by Year

Program Year	Accounts	Gas Savings (m^3)
2013	680	53,030,333
2014	573	46,195,015
2015	706	51,330,067

Enbridge RunitRight Participation Data

For RunitRight, the program tracking data includes projects claimed in the 2014-2015 program years. These projects were all implemented in 2013-2014; savings for a project in the program are claimed after one year of site metering is complete.

Table 12: Enbridge RunitRight Program Participation Metrics by Year Claimed

Program Year	Accounts	Gas Savings (m^3)
2014	45	625,088
2015	28	542,442

The RunitRight program has only one measure type. It also has several projects with negative savings. Negative savings (increases in energy use) are possible results from retro-commissioning projects, sometimes due to calculation method (billing analysis based savings without weather, occupancy adjustment or production adjustment) or due to actual increases in energy use. Negative saving measures need to be handled carefully in ratio estimation: high FR on large negative savings projects can result in overall program FR <0, which is not a valid FR result.[19] Our recommended approach to the problem is to produce and apply ratios with separate domains for positive and negative savings projects. Applying the ratios by separate domains based on positive or negative savings provides meaningful FR ratios and accurate net savings. Ratio estimation by domains is described in detail in Appendix E .

Union Custom Participant Data

The participant data files provided by Union included energy efficiency projects claimed during the 2013-2015 program years. The records in the tracking data are per installed measure, so there may be multiple rows per project if more than one measure is installed, and there may also be more than one project per account. There are 67 accounts that appear in both the 2015 program year and the 2013/14 program years.

Table 13: Union Participation Metrics by Year

Program Year	Accounts	Gas Savings (m^3)
2013	352	369,438,742
2014	392	285,752,549
2015	462	201,620,726

Task 2.2: Define the Unit of Analysis

Following data exploration, we defined the *unit of analysis*, which established the level at which data will be analyzed but not the level at which it will be collected, which is the *sampling unit*. We further discuss this distinction, how the sampling unit is defined and provide more detail on defining the unit of analysis in Appendix A .

The definition of the unit of analysis is one of the most important and least discussed aspects of DSM program evaluation. Consider the following four dimensions: end -use, measure type, equipment or “action,” and calculation approach.

- The **end use** can be important in decision making because lowering the cost per unit produced is a different decision than lowering the cost of heating a facility or office, for example. It can also be used as a proxy for the complexity of the calculation, as process-related end uses tend to have more complex and site-specific calculation approaches. End use can be used in surveys by listing the measure types that fall into the category; however, this is not ideal for NTG as the program’s influence on decision making may differ by measure type, affecting the attribution response.
- **Measure type** is important for surveys to aid participant recall by providing a concrete, simple description of what equipment was altered or installed. This aggregation is less appropriate for CPSV where the calculation method may differ.
- **Equipment or Action** is a very important distinction for NTG. Continuous improvement actions, such as maintenance, operations, and optimization, have fewer barriers to implementation than equipment purchases due to lower total cost, shorter term planning horizons and often fewer approvals. Businesses typically have separate budgets for capital and operating expenses. Purchases of new or replacement equipment falls under a capital budget, while actions are usually part of the operating budget or performed by salaried employees. Capital budgets typically have long term planning and allocation, while an operating budget is by nature more flexible to conditions in a given year. The ability of programs to affect equipment and action decision making is necessarily different as well. For the unit of analysis, actions were defined by three categories: maintenance, operational improvement, and optimization.

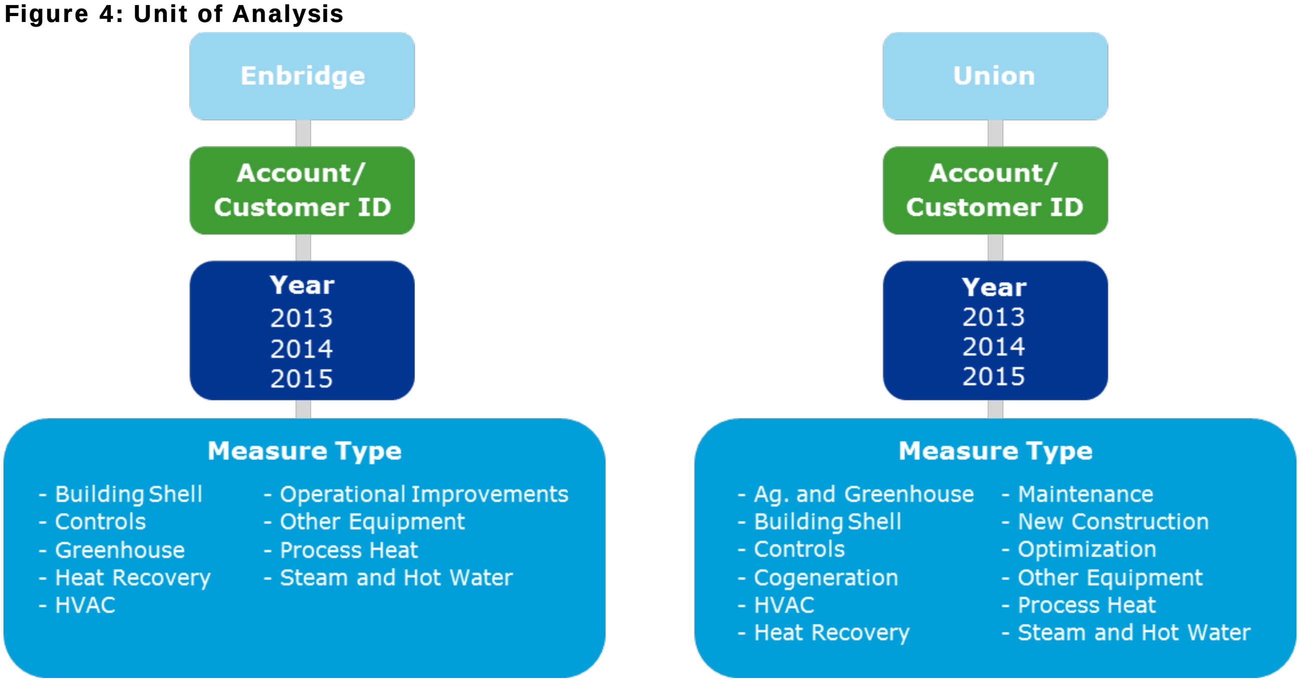
[19] Free ridership on negative savings results in more program savings, rather than less.

- **Calculation type** is important for CPSV. Simple, commonly implemented measures in custom programs do not require the same depth of data collection to verify calculations and inputs as more complex measures. Simple measures also use standardized calculation approaches that reduce variance. Evaluators tend to find fewer adjustments and, even when adjustments are found, the adjustment often affects all measures of a calculation approach similarly. For CPSV of the 2015 program year, the utilities indicated that projects that used eTools or standard spreadsheet calculation approaches could not be easily identified in the tracking data. In lieu of this information, evaluation engineers reviewed the provided program tracking database and categorized measures as simple or complex based on the best technology, enduse, and facility type descriptors available. A subset of measures (generally "other" measures) was categorized after the utilities provided additional detail.

Aggregating across any of the four listed dimensions is a trade-off of accuracy for increased precision, reduced customer burden and reduced evaluation costs. Not aggregating makes the same trade-off, but in reverse.

The unit of analysis for the evaluation, presented in Figure 4, aggregates the data to the utility, account, year, and measure type. For Union, aggregating the tracking data to the unit of analysis reduced the number of records from 744 to 597 records for 2015 and from 1,468 to 1,091 records for 2013 to 2014. For Enbridge, the number of records for 2015 decreased from 955 to 858 records and for 2013 to 2014 decreased from 1,648 to 1,511 records.

Figure 4: Unit of Analysis



Once aggregated to the unit of analysis, Union had an average of 1.5 units of analysis per account in 2013 and 2014 and 1.3 units per account in 2015[20] while Enbridge has an average of 1.2 units per account in 2013, 2014, and in 2015. In general, Union accounts tended to have more units of analysis per account than Enbridge accounts. Only 26 Union accounts have 5 units and none had more than 5. For Enbridge, 9

[20] We are assuming a 1:1 account to customer ratio for sampling. For the analysis, customer will be defined by contact information (phone number primarily), which is not included in the provided tracking data.

accounts have 4 units and no accounts have more than 4. This will facilitate data collection, since it's reasonable to ask about 3-4 units.

At this time we are unable to comment on the number of units per customer, because some customers will likely have multiple accounts. Customers will be defined by their contact information which will be requested along with the documentation request following submission of the scope of work.

Task 2.3: Stratify the NTG and CPSV Data

There is a balance between having too many and too few strata.[21] In sample designs, more strata allow the design to control representativeness and estimated precision along more dimensions. Having more strata does not hurt overall precision, but it can increase the sample sizes required. Each stratification level serves to improve efficiency, improve representativeness, or both.

There are four populations across which the evaluation findings will be completely separate from one another.[22] These populations are defined by having separate program designs. The divisions between these populations are hard lines; none of the reported ratio results will include a mix of information across these populations. We can think of this as four evaluations using a common methodology and data collection effort.

1. Union Large Volume
2. Union Custom C&I
3. Enbridge Custom C&I
4. Enbridge RunitRight

Within the stratification segments (see Table 10) we categorize measures to improve the efficiency and representativeness of the sample. [23] The stratification for the 2015 data collection effort balances the needs of two studies, with the CPSV sample a subset of the NTG sample. DNV GL's experience is that each has differing measure categorization priorities.[24]

- For NTG the measure categorization most predictive of free ridership rates is whether the project is installation of efficient equipment or whether the project was an action taken with existing equipment, regardless of whether that action is maintenance or an optimization that leads to energy savings.
- For CPSV the measure categorization most predictive of verification rates is a simple calculation versus one that is complex. Simple projects that follow consistent approaches and vary less from site to site typically have verification rates with lower variance than more complex projects that require more site specific knowledge and truly custom calculations. Stratifying by rigour allows us to assign a lower ER (0.3) to the simple project strata and higher to the more complex strata (0.4 ER) which provides better sample allocation. Simple strata projects will receive a TSER verification, while complex strata projects will receive an onsite verification.

---

[21] DNV GL agrees with the approaches described in "Sampling Methodology for Custom C&I Programs" which was prepared by Navigant for the TEC in 2012 and used to inform previous CPSV sample designs. Our sample design approach is consistent with the approaches recommended and follows the recommended seven step process (pages 17-23).
Dan Violette, Ph.D. & Brad Rogers, M.S., MBA, Navigant Consulting, Inc. "A Sampling Methodology for Custom C&I Programs," Prepared for: Sub-Committee of the Technical Evaluation Committee. November 12, 2012 (Revised October 28, 2014).

[22] For the CPSV, LI MF will be reported with MR MF either together with Custom C&I or as a separate Multi-Family domain, depending on final sample sizes and precisions.

[23] Page 14 in the Navigant report provides an explanation of the rationale for stratification.

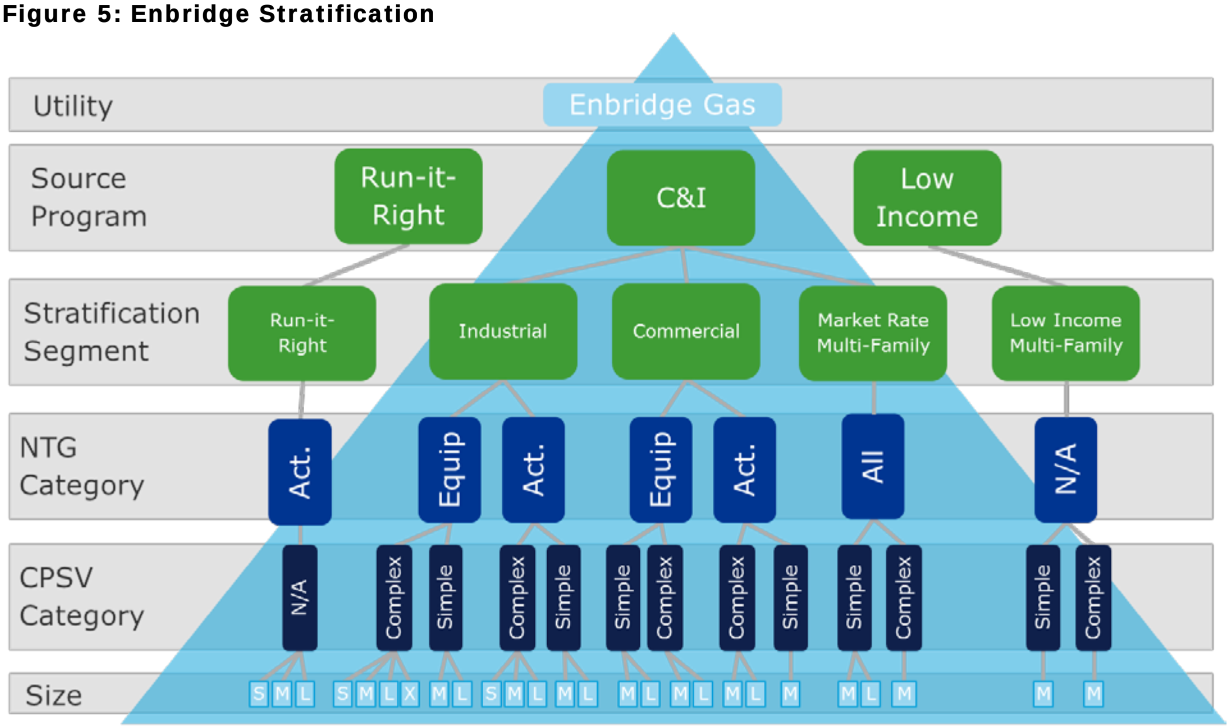
[24] The current stratification plan has more aggregated program segment categories than were described in the original proposal. When developing the proposal sample design we did not have access to the data or savings amounts specific to measure types.

The final stratification level segments projects by the magnitude of energy savings resulting from that project. Large projects represent a greater portion of the population, so sampling them at increased rates will result in greater precision with fewer verification visits or calls. Smaller projects must also be sampled to ensure representativeness. In the final 2015 sample design, DNV GL used cumulative savings as a measure of size; for 2013/14 sampling annual savings were used. Cumulative gross savings were not provided for some of the 2013/14 programs.[25] In terms of sample allocation, using cumulative savings selects longer life measures at a higher rate than would occur if annual savings were used.

It is important to note that the stratification used for sampling and expansion does not need to correspond directly to the level of reporting. For example, while we have chosen to use broad categories of customer segments in our stratification, this does not preclude reporting or applying ratios by more disaggregate customer segments. Our intended (pending final precisions) application domains are provided later in this section.

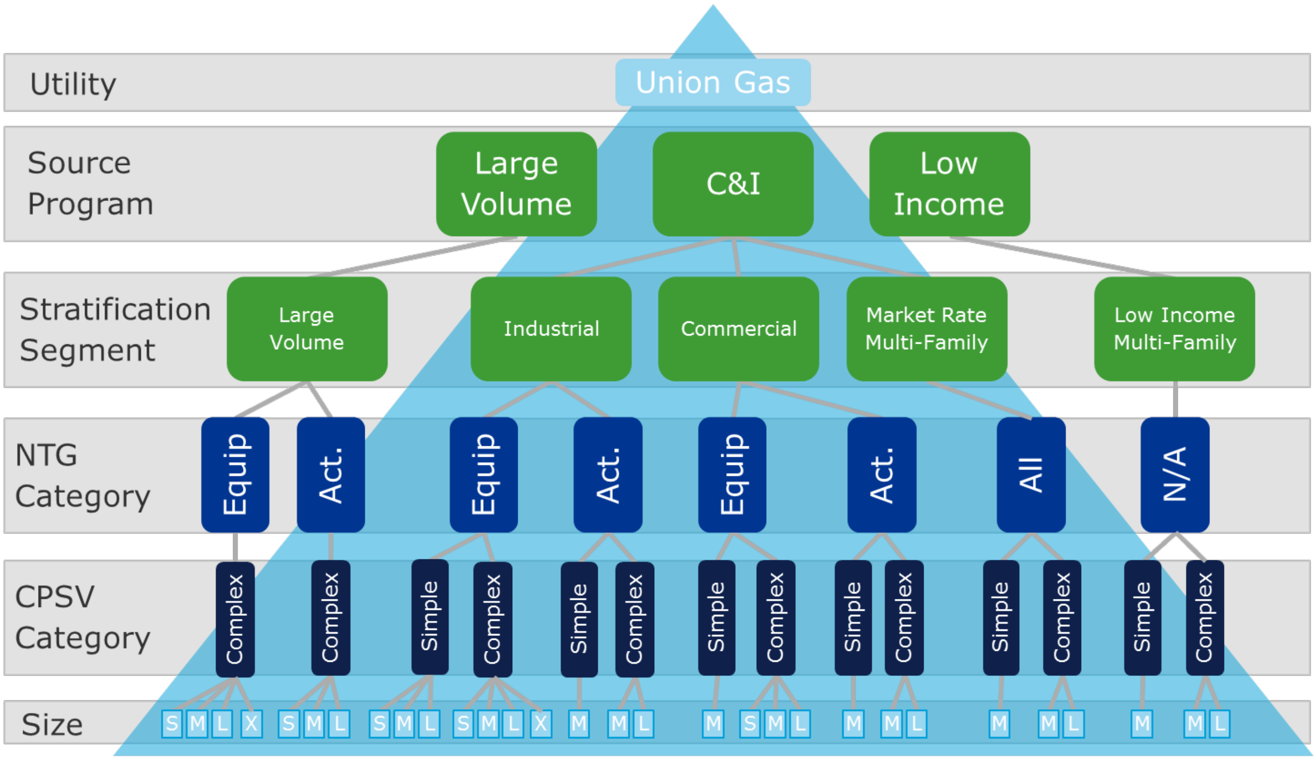
Figure 5 and Figure 6 show the CPSV and NTG stratification for Enbridge and Union respectively.

Figure 5: Enbridge Stratification



25 The August 4, 2016 data provided by Enbridge included cumulative savings for 2015, but not 2013 and 2014.

Figure 6: Union Stratification



Task 2.3: Design the Samples

Critical to the sample size determination is the error ratio for each sampling cell with respect to the ratio to be estimated. The error ratio for ratio estimation is the equivalent of the coefficient of variation for estimation of a population mean. A higher ER assumption results in a larger required sample size.

The error ratios used in the sample design are lower than typical ER assumptions[26] due to the stratification described above:

- 0.6 for FR
- 0.3 for Simple CPSV strata (TSERs)
- 0.4 for Complex CPSV strata (Onsites)
- 0.35 for Complex Multi-Family CPSV strata (includes both MR MF and LI MF)

2015 Enbridge Stratification

The 2015 Enbridge stratification is presented in Figure 7. The final stratification includes 4 evaluation programs, two NTG categories, two CPSV categories and up to three size categories optimized for sampling efficiency. In total there are 26 Strata.

Table 14 shows the 2015 Enbridge sample design in table form with the expected relative precisions and sample for targeted programs shown for each portion of the study. Data collection for FR will be completed

[26] In general, a typical ER for FR with size-only stratification would be 0.7, while for CPSV, 0.4 or 0.5 would be appropriate.

through IDI, while the CPSV group indicates strata that will have onsite data collection (complex) or TSER (simple).

Table 14: Enbridge CPSV and FR Sample Design

Stratification Segment	NTG Group	CPSV Group	Size Strata	Sample Frame		FR Sample Design		CPSV Sample Design	
				N	m3	n	Rel. Prec.	n	Rel. Prec.
Industrial	Action	Complex	3	13	18,898,127	8	10%	7	10%
		Simple	2	8	4,964,165	4		4	
	Equipment	Complex	4	70	276,569,945	24		15	
		Simple	2	25	43,925,065	6		6	
Commercial	Action	Complex	2	3	10,988,780	3	10%	3	10%
		Simple	1	24	3,875,430	4		4	
	Equipment	Complex	3	59	61,573,901	22		9	
		Simple	2	293	236,656,958	34		10	
Market Rate Multi-Family	All	Complex	1	53	23,584,650	8		5	
		Simple	2	175	129,568,929	19		8	
Low Income	All	Complex	1	6	5,125,020	0	N/A	2	
		Simple	1	104	58,676,555	0		6	
RunitRight	Optimization	IDI	3	28	2,712,210	17	10%	0	N/A
Total				**861**		**149**		**79**	

Table 15 shows the anticipated relative precisions for less aggregated program segments. We expect that the final relative precisions will be close to 90/10 for these segments as well as the targeted programs above. The domains that will be included in the report and the domains which will be used for ratio application will be determined based on the decision making structure provided in Appendix E.

Table 15: Enbridge Expected Precisions by Program Segment

Stratification Segment	Sample Frame		NTG		CPSV	
	N	Savings	n	Rel Prec.	n	Rel Prec.
Industrial	116	344,357,302	42	10%	32	10%
Commercial	379	313,095,069	63	11%	26	14%
MR MF + LI MF	338	216,955,154			21	14%
MR MF	228	153,153,579	27	21%		
RunitRight	28	2,712,210	17	10%		

2015 Union Stratification

The Union stratification is shown in Figure 8. The final stratification includes 4 evaluation programs, two NTG categories, two CPSV categories and up to three size categories optimized for sampling efficiency.[27] In total there are 30 strata.

Table 16 shows the 2015 Union sample design in table form with the expected relative precisions and sample for targeted programs shown for each portion of the study. Data collection for FR will all be completed through IDI, while the CPSV group indicates strata that will be have onsite data collection (complex) or TSER (simple) for CPSV.

[27] Size strata have different cutoffs between large/medium/small etc within each category. The cut points were determined using a Model Based Stratified Sampling (MBSS) algorithm that optimizes the cut-points and sample allocation to produce the best precision overall for a given number of sampled units.

Table 16: 2015 Union CPSV and FR Sample Design

Stratification Segment	NTG Group	CPSV Group	Size Strata	Sample Frame		FR Sample Design		CPSV Sample Design	
				N	m3	n	Rel. Prec.	n	Rel. Prec.
Industrial	Action	Complex	1	21	75,487,148	7	11%	6	10%
		Simple	1	44	102,200,503	4		3	
	Equipment	Complex	4	136	862,582,429	35		17	
		Simple	3	111	165,066,284	10		9	
Commercial	Action	Complex	2	8	81,635,903	5	10%	4	
		Simple	1	13	22,029,892	6		3	
	Equipment	Complex	3	109	142,631,725	14		7	
		Simple	1	42	14,831,059	5		3	
	All	Complex	2	6	7,409,515	3		2	
		Simple	1	1	44,260	1		1	
Low Income	All	Complex	2	2	1,454,295	0	N/A	2	
		Simple	1	35	4,466,365	0		3	
Large Volume	Action	Onsite	3	35	404,398,149	10	10%	8	10%
	Equipment	Onsite	4	37	846,481,549	22		13	
Total				**579**		**115**		**75**	

Table 17 shows the anticipated relative precisions for less aggregated program segments. We expect that the final relative precisions will be close to 90/10 for these segments as well as the targeted programs above. The domains that will be included in the report and the domains which will be used for ratio application will be determined based on the decision making structure provided in Appendix E.

Table 17: 2015 Anticipated Precisions by Program Segment

Stratification Segment	Sample Frame		NTG		CPSV	
	N	Savings	n	Rel Prec.	n	Rel Prec.
Industrial	312	1,205,336,364	56	11%	35	12%
Commercial	172	261,128,579	30	12%	17	16%
Large Volume	72	1,250,879,698	32	10%	21	10%
MR MF + LI MF	44	13,374,435			8	10%
MR MF	7	7,453,775	4	25%		

Task 2.4: Spillover Sample

The sample design for spillover omits the CPSV category, but is otherwise consistent with the sample design for the 2015 FR and CPSV evaluation task. For spillover the ER used was 0.8; 90/10 precision was targeted. The assumed ER of 0.8 reflects the weaker correlation between SO and the size of the original measure than there is for FR or CPSV.

2013/14 Enbridge Stratification

The 2013/14 Enbridge stratification is presented in Figure 7. The final stratification includes 4 evaluation programs, two NTG categories, and up to six size categories optimized for sampling efficiency. In total there are 28 strata.

Figure 7: 2013/14 Enbridge Stratification

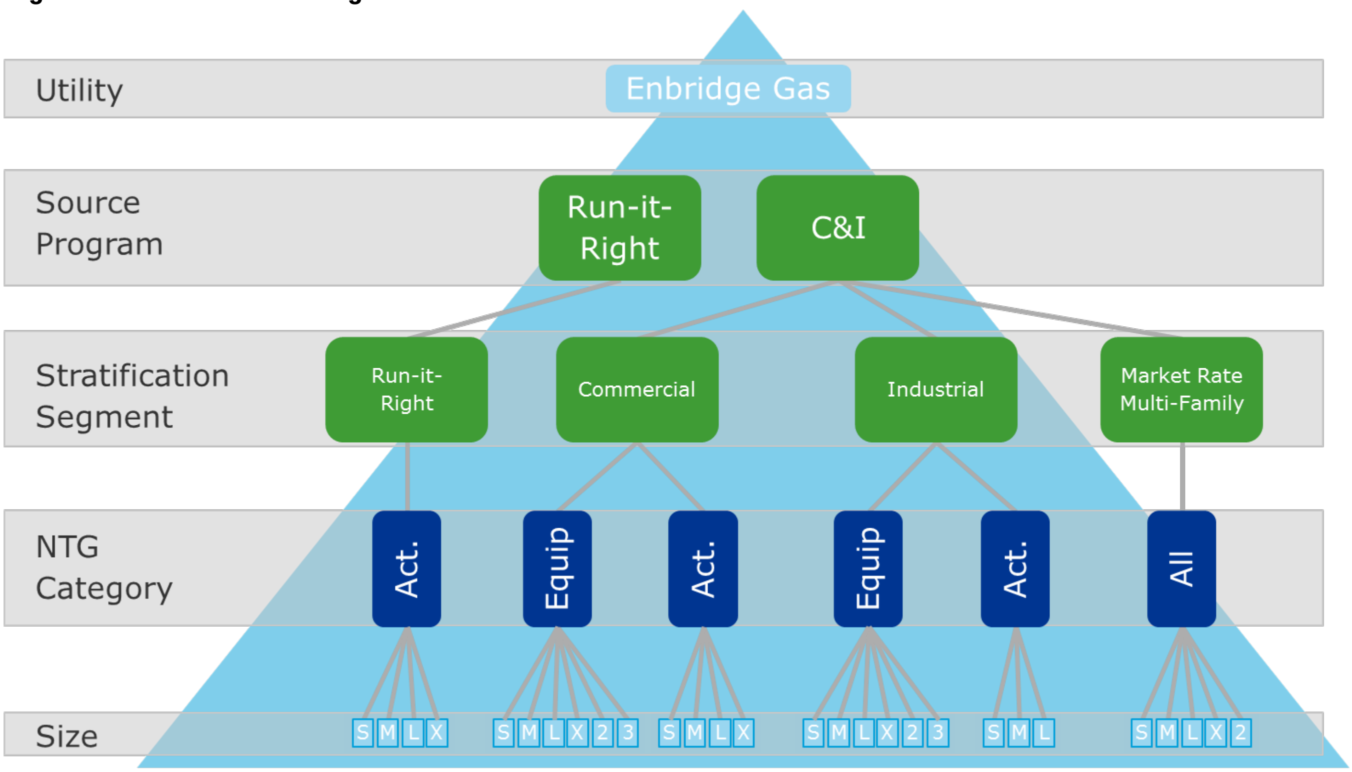


Table 18 shows the number of accounts and savings in the sample frame and the targeted spillover sample size for each grouping. The domains that will be included in the report and the domains which will be used for ratio application will be determined based on the decision making structure provided in Appendix E.

Table 18: 2013/14 Enbridge Spillover Sample Design

Stratification Segment	NTG Group	Size Strata	Sample Frame N	Sample Frame m3	SO Sample Design n	SO Sample Design Rel. Prec.
Industrial	Action	3	40	5,067,923	20	10%
	Equipment	6	191	41,899,589	50	
Commercial	Action	4	79	4,604,864	25	10%
	Equipment	6	603	27,240,429	60	
MR MF	All	5	553	20,412,543	65	
RunitRight	Action	4	45	625,088	26	10%
Total			**1,511**		**246**	

2013/14 Union Stratification

The Union stratification is presented in Figure 8. In total there are 35 strata.

Figure 8: 2013/14 Union Stratification

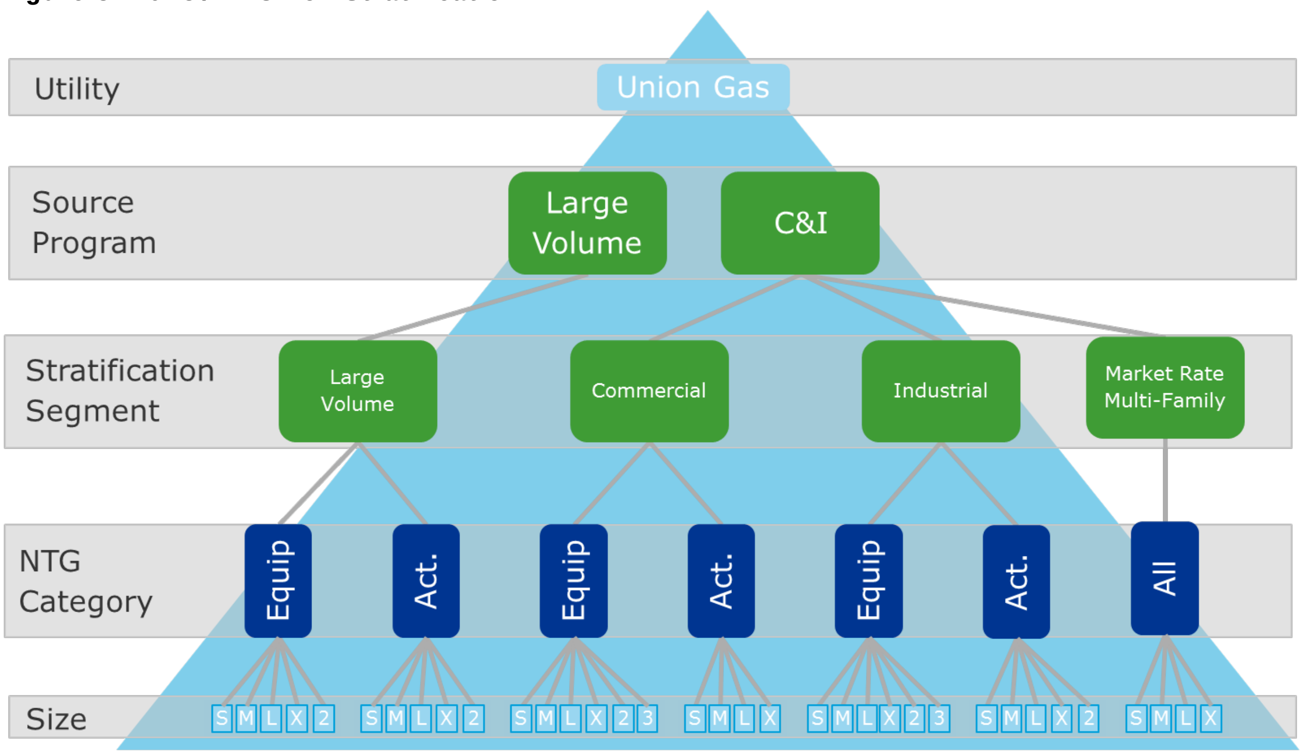


Table 19 shows the number of accounts and savings in the sample frame and the targeted spillover sample size for each grouping. The domains that will be included in the report and the domains which will be used for ratio application will be determined based on the decision making structure provided in Appendix E.

Table 19: Union Spillover Sample Design

Stratification Segment	NTG Group	Size Strata	Sample Frame N	Sample Frame m3	SO Sample Design n	SO Sample Design Rel. Prec.
Industrial	Action	5	167	64,448,800	38	10%
	Equipment	6	412	107,347,726	57	
Commercial	Action	4	74	9,687,715	24	10%
	Equipment	5	190	15,744,760	40	
MR MF	All	2	38	564,428	8	
Large Volume	Action	5	130	317,638,812	38	10%
	Equipment	5	94	139,759,050	33	
Total			**1,105**		**238**	

Task 2.5: Prepare the Sample and Backup Sample

Once we have completed the final sample design, we will submit a data request to the utilities. For the 2015 NTG sample we will request documentation and contact information for 50 percent more projects that are in the primary sample. The 50 percent additional constitutes the initial backup for the FR sample. This corresponds to a minimum 66 percent response rate. If response rates are lower than 66 percent in specific strata, we will request documentation and contact information for additional projects in the stratum as needed to meet targets.

For the 2013/14 spillover sample we will request contact information and a review of the measure description to be used in the survey for three times the number of sampled projects. These additional projects are the initial backup for the spillover sample. The number requested corresponds to a minimum 33 percent response rate. If response rates are lower for the SO CATI than 33 percent in specific strata, we will request contact information for additional sites as needed.

We will not request project documentation for the entire spillover sample. Once we have identified the sites that require follow up engineering interviews we will request documentation for necessary sites and up to 20 additional sites beyond those that will receive follow up calls Based on prior experience we anticipate conducting follow up calls for approximately 10 percent of the original sample. Overall this staging of requests will reduce the amount of project documentation that the utilities need to provide, while ensuring efficient data collection and respondent confidentiality.

Backups for each sample will only be contacted if needed to meet targeted number of completes.

Once we have received the requested contact information, we will identify instances where a contact was involved in multiple projects, even across sites. While the projects are conducted at the site level, the decision maker, technical expert, or vendor may have been involved in projects at multiple sites. For example, multiple participating sites for the same retail chain may have one energy manager from the corporate office but the technical expert may be site specific. Using this contact information and taking into account cross-site involvement, we will assemble the CATI and the IDI sample frame.

Task 3: Data Collection

Data collection for the program includes interviews with program managers, staff and energy advisors; CATI surveys and IDIs with program participants; on site verification at participating customer sites; and IDIs with participating vendors. Interviews with program staff and energy advisors are for informational purposes only. CPSV, FR and SO results will be based on data collected directly from participating customers and vendors.

Table 20: Task 3 Subtasks and Completion Status

Key Tasks
☒ Task 1: Project Kickoff
☐ Task 2: Sample Design
☐ **Task 3: Data Collection**
☒ **Task 3.1: Program Managers and Staff Interviews**
☒ Union Gas Program Portfolio Management
☒ Enbridge Commercial Program Managers
☒ Enbridge Industrial Program Managers
☒ Enbridge RunitRight Program Managers
☐ **Task 3.2: Program Energy Advisor Interviews**
☐ Energy Advisor Interview Guide
☐ Up to 10 pre-survey interviews
☐ **Task 3.3: Program Participants**
☐ SO CATI survey Instrument
☐ Up to 502 CATI surveys completed
☐ FR IDI Instrument,
☐ Up to 280 FR IDIs completed
☐ CPSV Data Collection forms
☐ Up to 107 site visits completed
☐ Up to 62 TSERs completed
☐ Approx. 50 SO follow up IDIs completed
☐ **Task 3.4: Participating Vendors**
☐ In-Depth Interview Instrument
☐ Up to 80 IDIs completed
☐ Task 4: Data Analysis
☐ Task 5: Reporting
☐ Task 6: Project Management

Objectives

The objective of the data collection task is to collect

- Program manager, staff and energy advisor information on program services and influence on participants and vendors to inform other data collection efforts
- Participant information on timing, efficiency, and quantity to inform FR analysis
- Vendor information on timing, efficiency, and quantity to inform the FR analysis
- On-site and telephone data from participants about equipment and operations to inform the CPSV
- Participant information on energy saving projects undertaken outside of the program(s) and program influence thereof to inform the spillover analysis

Activities

Each of the data collection activities supports either understanding program influence on energy savings projects or verifying gross energy savings.

1. Group interviews with utility program managers, and staff informed the development of the scope of work. These interviews focused on how the programs are designed and operate to influence projects

directly and indirectly through incentives, technical assistance, vendor relationships and long term customer management.

2. Program energy advisor interviews will be scheduled after submission of the draft SOW. These interviews will focus on the specifics of program interactions with customers. The intent of the interviews is to ensure that the FR framing in the IDIs and CATI covers the range of program activities that may have influenced decisions to implement projects.

3. Program Participants are the primary source of data for the evaluation and verification.

    a. Spillover CATI surveys will be conducted to identify customers with potential spillover projects. The CATI will utilize the FR framing to aid customer recall of the original program measure and the program interactions associated with it. Then the CATI will ask the customer if any other energy saving actions have been taken since the original measure and, if so, whether these actions were influenced by the prior program participation.

    b. FR IDIs will be conducted to estimate the free ridership for the 2015 program. These IDIs will ask primarily open ended questions about program and other influences in a FR framing section and then will ask a series of questions to estimate free ridership for each measure. A subset of these IDIs will include gross savings verification questions (for the TSER sample) and a subset will also be asked SO questions (if they are also in the SO sample for another measure).

    c. On site visits will collect data to support verification of gross savings estimates (onsite sample). These visits will occur after the initial FR IDI for sites in the onsite sample.

    d. Engineering follow up IDIs will collect data to support spillover savings estimates and free ridership of the original program measure (where applicable)

4. Participating vendors will provide supplemental data for FR estimates for customers who indicate vendor influence on their decision to implement program measures.

Table 21 is a summary of the targeted completes by data collection type. For spillover the sample design targets 484 completed surveys. A portion of these surveys will be completed as a spillover module addressing 2013/14 projects at the end of the FR interview with 2015 participants who also participated in 2013/14. A more detailed breakdown of our target number of surveys and interviews is provided in the description of the methodology in Task 2: Sample Design.

Table 21: Target Number of Completed Surveys/Interviews

Target Group		Number of Interviews/Surveys		
		Enbridge	Union	Total
Program Manager Interviews		3	1	4
Energy Advisor Interviews		≤10	≤10	≤20
Spillover Only	Participant CATI Surveys	≤246	≤238	≤484
	Participant Follow up Interview	~25	~25	~50
FR/CPSV/Spillover	Total Participant IDIs	151	121	272
CPSV	Participant Site Visits	40	57	97
	Participant TSERs	38	22	60
FR	No CPSV	73	42	115
FR/Spillover	Participating Vendors	~30	~30	≤62

Shortly after this scope of work is provided to the EAC for review, DNV GL will contact the utilities to schedule interviews with program EAs to support prompt submittal of the draft interview guides and CATI surveys to the EAC for review.

A comment matrix will describe how comments or suggestions from reviewers were addressed in the final interview guides and survey instruments. While this review process is ongoing, we will also request contact information from the utilities to ensure that we are talking to the appropriate people, have the necessary contact information, and work with the utilities to prepare draft advance letters.

Deliverables

- Program participant Spillover CATI survey instrument (draft and final), including proposed scoring algorithm memo
- Program participant Free ridership IDI guide (draft and final), including proposed scoring algorithm memo
- Participating vendor IDI guide (draft and final), including proposed scoring algorithm memo
- Program manager and staff IDI guide (draft and final)
- Program Energy Advisor IDI guide (draft and final)
- CATI and IDI participation advance email & mail scripts

Task 3.1: Program Managers/Staff In-Depth Interviews

In order to better understand program logic, methods, execution, and intent, DNV GL conducted IDIs with program managers and then program staff. The interviews informed the development of the scope of work in the planning of data collection and analysis. These interviews focused on:

- Understanding how the program is designed
- Understanding how the program is implemented and marketed
- Understanding the program theory and logic
- Identifying key staff such as Energy Advisors and what roles they play

- Identifying how staff understand decisions are made by customers
- Identifying how communication among customers, program staff and vendors parties occurs.

DNV GL staff interviewed program staff from Enbridge and Union on the following dates:

Table 22: Program Manager Interviews

Date	Company	Program
1/22/2016	Union Gas	Program Management - Portfolio
1/25/2016	Enbridge	Commercial Programs Interview
1/29/2016	Enbridge	Industrial Programs Interview
1/29/2016	Enbridge	RunitRight Program Interview

Task 3.2: Program Energy Advisors

The evaluation will request interviews with ten Energy Advisors prior to submitting the final program participant survey/interview instruments, in order to better inform those instruments. Five will be interviewed from each of Enbridge and Union Gas (10 total interviews). For these initial, non-project specific interviews, we will ask the utilities to select the Energy Advisors who they feel will be most helpful to the evaluation in terms of how the program influences projects and works with vendors. These interviews will inform our participant data collection guides to ensure that they address the actions of each of the programs that this evaluation is addressing. The outline for the initial Energy Advisor interview is provided in Appendix F .

The evaluation will attempt to schedule an additional 10 energy advisor interviews will be scheduled prior to fielding the participant IDIs. Five will be interviewed from each of Enbridge and Union Gas (10 total interviews). The energy advisors will be those with the five largest projects in the Union and Enbridge programs respectively. We will discuss the two largest projects in the sample that are associated with each energy advisor selected. These interviews will consist of talking through the FR framing topics with the energy advisor regarding each project. Following the interviews project specific probes will be added to the specific project's interview guide as necessary. Added probes, with participant identifying information redacted, will be provided to the EAC prior to administering the participant IDI for transparent review to ensure that any probes added are "non-leading" and will not bias the FR results.

The outline for the project specific energy advisor interviews will parallel the FR Framing module in the participant IDI.

Task 3.3: Participant Data Collection

Participant data collection will be a combination of CATI surveys, in-depth-interviews and onsite visits. The combination of data collection approaches that will be used for a specific customer depends on which samples the customer has been selected for.

Figure 9 shows the populations and samples that a single customer may a part of. The figure shows how a customer may be included in as many as three samples or as few as none. The CPSV Onsite and TSER samples are mutually exclusive and a subset of the FR Sample. The NTG-Only sample is the portion of the FR sample that was not selected for the CPSV. The Spillover sample overlaps each of the Onsite, TSER and NTG-Only samples and includes many customers that are included in the 2015 samples.

Figure 9: Population and sample overlap

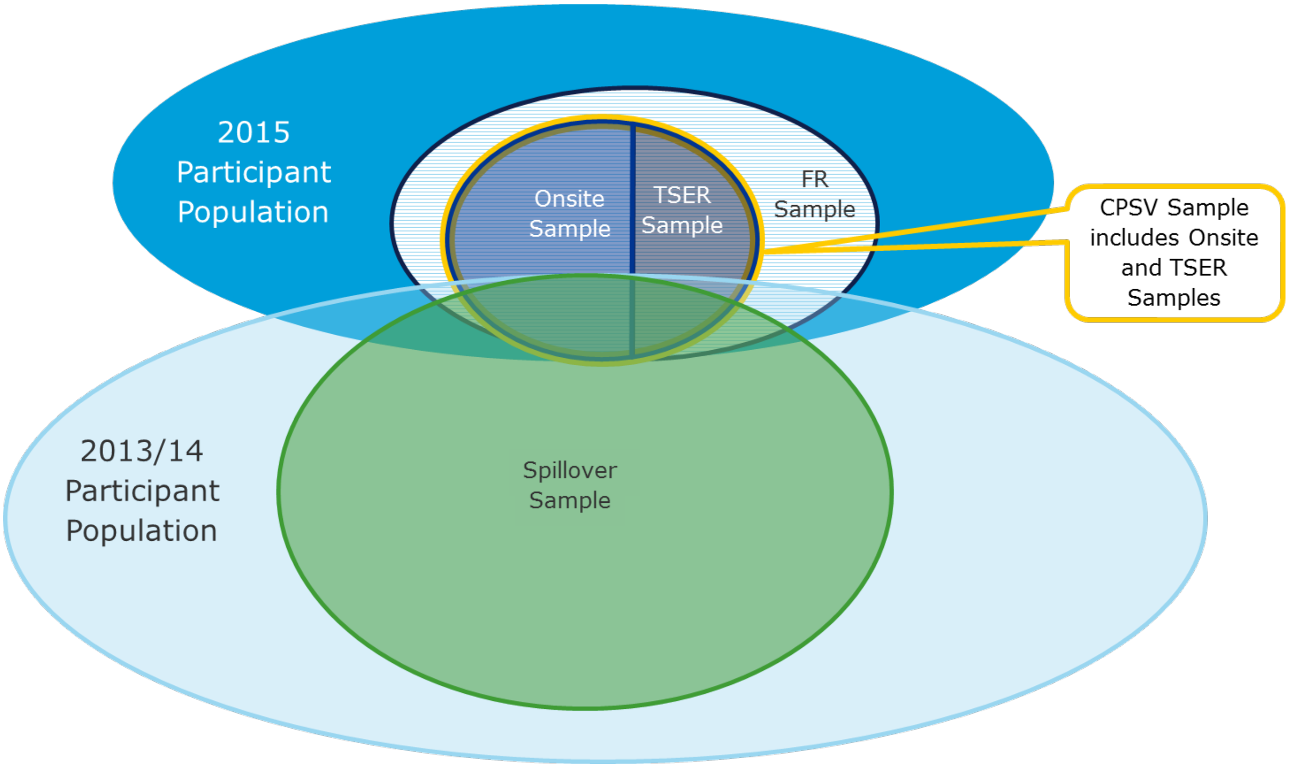


Table 23 shows the data collection efforts that will be attempted with each sample group.

- Advance letters will be sent to all of the customers selected for each sample and backup sample.
- CATI surveys will be conducted with customers sampled for spillover only.
- FR focused IDIs will be administered for any customer selected in the FR sample (made up of the Onsite, TSER and NTG-only samples).
- TSER interview modules will be included in the FR IDI for the TSER sample.
- Following the initial IDI, Onsite sample customers will receive a scheduling call to schedule an onsite visit.
- Spillover sample customers will receive a follow up IDI if the initial survey or interview indicates the potential for spillover at one of the customer's sites.
- Vendors who worked with customers in the NTG sample will receive a project specific vendor NTG interview for projects where the customer indicates vendor influence.

Table 23: Data Collection for each Sample Group

2015 Participants	2013/14 Participants	Target Completes			Advance Letter	CATI	IDI	Onsite Sched & Visit	SO Follow up	Vendor Interview*
		Enbridge	Union	Total						
Onsite Sample	Spillover Sample	23	19	10	✓		✓	✓	✓	✓
	Not Sampled/Non-Part	17	38	97	✓		✓	✓		✓
TSER Sample	Spillover Sample	30	13	2	✓		✓		✓	✓
	Not Sampled/Non-Part	8	9	60	✓		✓			✓
NTG-Only Sample	Spillover Sample	64	15	2	✓		✓		✓	✓
	Not Sampled/Non-Part	9	27	109	✓		✓			✓
Not Sampled or Did Not Participate	Spillover Sample	≤246	≤238	≤484	✓	✓			✓	

*Vendor interviews are with participating vendors, not participating customers.

Figure 10 shows the data collection flow and assignments for the CPSV and FR sample projects. Primary responsibility for each task is colour coded by company in the shape “fill.” Outlines of each shape indicate the company with secondary responsibility.

- Advance letters will be coordinated and sent to the sample population by DNV GL and the Gas Utilities.
- FR IDIs will be conducted by a qualified DNV GL or Itron interviewer for all customers selected in the Onsite and NTG-only samples.
- FR+TSER modules will be included in the TSER sample interviews and will be conducted by DNV GL evaluation engineers with experience administering NTG IDIs.
- Onsite sample customers will receive a scheduling call from a Stantec recruiter to schedule an onsite visit. Gas utilities will be asked to facilitate scheduling on a case-by-case basis. In order to avoid the appearance of bias, the gas utilities will not be asked to take the lead on scheduling onsite visits.
- Onsite verification will be carried out by qualified Stantec engineers. Depending on need, some of the most complex projects may have onsite verification completed by a DNV GL or Itron engineer. Gas utility representatives will be encouraged to facilitate and observe the onsite portion of the verification.
- Vendor IDIs will be administered by DNV GL or Itron interviewers for applicable projects.

Figure 10: Data collection flow for CPSV and FR Sample Projects

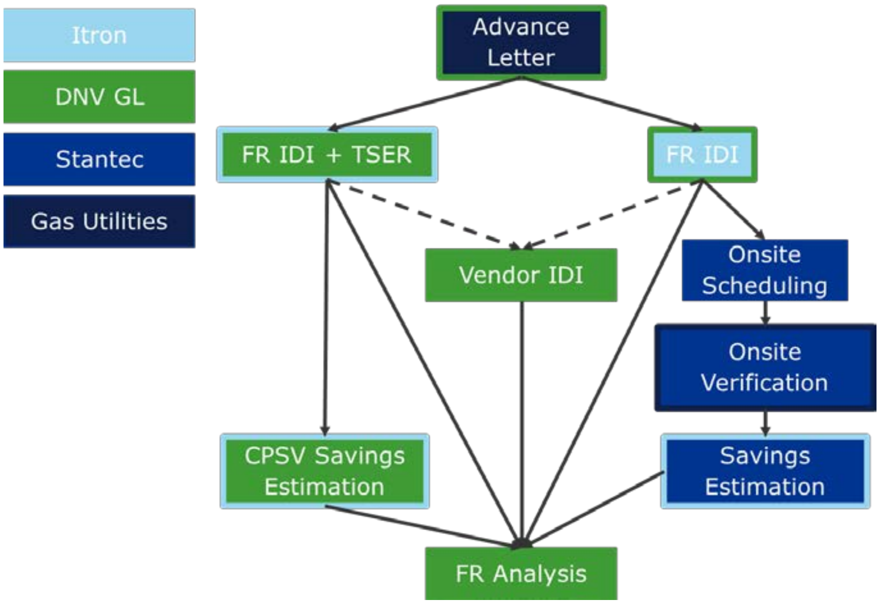
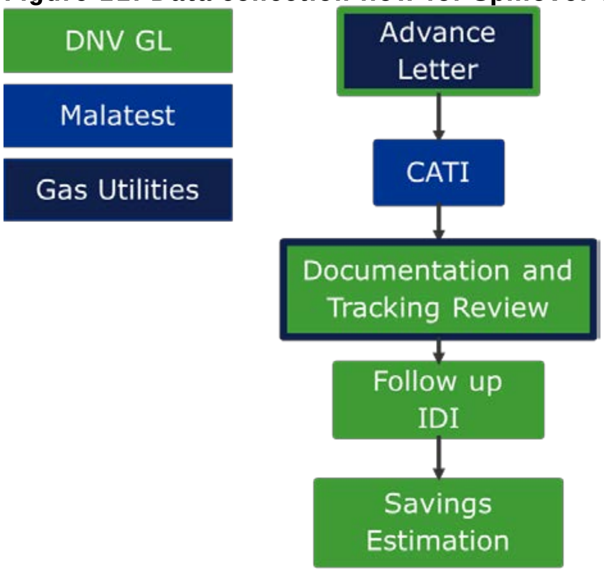


Figure 11 shows the data collection flow for the spillover sample projects. Primary responsibility for each task is colour coded by company in the shape "fill." Outlines of each shape indicate the company with secondary responsibility.

- Advance letters will be coordinated and sent to the sample population by DNV GL and the gas utilities.
- CATI Surveys will be administered by Malatest under direction from DNV GL. Not shown in the figure, a small number of customers selected for both the spillover and FR sample will have spillover modules administered as part of their FR IDI in lieu of the CATI.
- DNV GL will request project documentation and non-custom program tracking data for customers who report potential spillover projects in the CATI survey (plus additional customers in order to preserve respondent confidentiality)
- After reviewing project documentation, A DNV GL Engineer will conduct an IDI with customers who report potential spillover in the CATI survey.

Figure 11: Data collection flow for Spillover Sample Projects



As Table 23 indicates the success of the project will depend on having flexible instruments with different modules depending on the sample(s) that the customer is selected for. Table 24 shows how these modules will be distributed across the sampled customer types. Each module includes a framing section to aid customer recall and a section that will be scored. The determination of vendor influence will be done based on questions that are part of the framing in the FR module.

Table 24: Data Collection Modules by Sample Groups for Initial Customer Contact

2015 Participants	2013/14 Participants	Initial Modules: Informed Respondent	CPSV	FR	SO
Onsite Sample	Spillover Sample	✓		✓	✓
	Not Sampled/Non-Part	✓		✓	
TSER Sample	Spillover Sample	✓	✓	✓	✓
	Not Sampled/Non-Part	✓	✓	✓	
NTG-Only Sample	Spillover Sample	✓		✓	✓
	Not Sampled/Non-Part	✓		✓	
Not Sampled or Did Not Participate	Spillover Sample	✓			✓
	Not Sampled/Non-Part				

There is no spillover module for the 2015 projects because not enough time has elapsed for the large majority of these participants to have done a spillover project. It is possible that some customers may have done a spillover project in this short timeframe. As the table indicates, we will not be collecting free ridership information from the SO-only participants as part of the initial CATI or IDI. This information is required for our participant spillover methodology, but only for projects that have associated spillover. To reduce customer burden for the majority of customers, we will collect these data as part of the follow up spillover interviews.

Conducting IDIs of customers with large or complex projects is a standard method for DNV GL and Itron, with experienced and expert interviewers conducting all interviews.[28] These interviews are conducted with the 'decision maker' – an informed respondent who has at least some say in whether or not to proceed with a project and is aware of the project's impacts.

DNV GL and Itron staff will conduct IDIs with customers in the FR sample. FR for each project is certain to have an effect on the final net savings. The outline for Participant IDIs is provided in Appendix F .

CPSV Data collection

Data collection for the TSER sample will be completed via the IDI as described above. Prior to the TSER IDI the interviewing engineer will review project documents and calculations to identify the specific CPSV questions to include in the interview. Following the interview the engineer will complete the TSER verification report, embedded below. Verification reports completed by DNV GL engineers will be reviewed by an Itron engineer and verification reports completed by Itron engineers will be reviewed by a DNV GL engineer.

Onsite sample customers will not have engineering questions asked during the IDI. Instead these customers will be asked permission for a follow up site visit. Customers who agree to the site visit will receive a follow up call from Stantec to schedule the visit. Utility staff will be informed of the scheduled visit and invited to attend. Following the onsite visit, the Stantec engineer will complete the onsite verification report, embedded below. An Itron engineer will review the report.

Appendix F has the template forms that each of the data collection approaches will use for the CPSV.

Completed verification reports will be compiled into a draft report to be reviewed by the OEB and EAC. The steps in the CPSV review process are shown in Table 25.

[28] Names and CVs of specific interviewers and engineers will be provided after the SOW has been approved and the data collection schedule is more certain.

Table 25: CPSV Steps

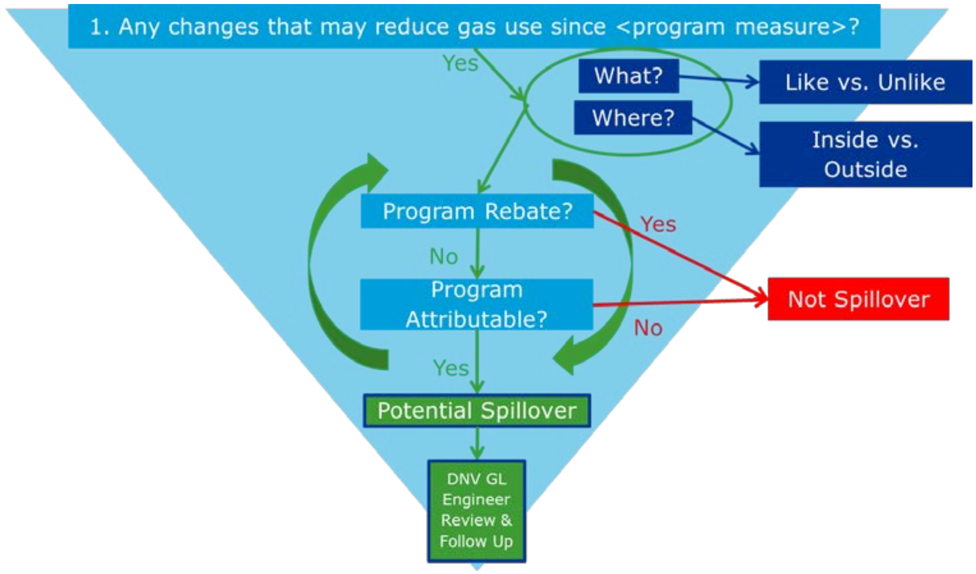
Step	Activity
1	NTG/CPSV Evaluation Team reviews project files provided by utilities ▪ Missing or incomplete documentation will be requested from utilities following review (final opportunity for utilities to provide new information).
2	NTG/CPSV Evaluation Team conducts IDI with customers ▪ Collects required CPSV data for TSER sample projects
3	NTG/CPSV Evaluation Team schedules site visits with onsite sample customers, informs utility
4	NTG/CPSV Evaluation Team conducts customer site visit ▪ Collects required CPSV data for Onsite sample projects
5	NTG/CPSV Evaluation Team drafts project verification reports ▪ Contacts utility staff/customer to clarify any site/operational details if needed.
6	EC Team conducts internal review of individual project verification reports ▪ Itron reviews projects verified by Stantec and DNV GL ▪ DNV GL reviews projects verified by Itron
7	EC Team shares draft report, including all site verification reports, with OEB for quality control, redacted as necessary.
8	EC Team (OEB team) shares final draft report with EAC, redacted as necessary
9	EAC provides written comments on final draft report
10	EC Team/OEB hold EAC meeting to discuss comments
11	EC Team finalizes report

Spillover Sample

Two of the challenges that SO presents for data collection are that many projects result in no spillover,[29] so evaluations need to contact a larger sample to achieve the desired 90/10 precision; and spillover can be hard to quantify since the program does not have the project documentation to calculate savings. To solve the first challenge, a large sample, we will use a CATI survey as our initial pass at identifying program spillover for most of the spillover sample. This will allow us to cast a wide net in our initial survey, contacting more customers. We will also be using a question sequence on these calls that seeks to start as wide as possible before narrowing down the potential spillover. Follow up IDIs will allow us to leverage evaluation engineer expertise to collect the data needed for savings estimates and to collect free ridership data for the original program measure. Figure 12 shows the high level approach that we will take in the CATI, with specific details provided in the spillover survey module and methodology memo.

[29] Unless the program is specifically designed to induce spillover.

Figure 12: High level approach to identifying potential spillover



Engineering Follow-Up Interviews for Spillover

For some projects, it will be necessary to follow up with an additional individual or individuals, aside from the 'decision maker'. Engineering follow up calls are a specialized form of IDI that are conducted between a DNV GL engineer and an individual at the customer site that can speak to the specific engineering specifications of the equipment. DNV GL will ask specific questions that will allow for the calculation of energy savings.

These interviews will be individually tailored, depending on equipment installations, with the goal of gaining information to calculate energy savings.

For like spillover measures and the subset of unlike spillover measures to which it applies,[30] the spillover follow up interview will also include the FR module for the original program measure.

Task 3.4: Participating Vendors (In-Depth Interviews)

Vendors that worked with customers on sampled projects will be interviewed if the participant indicates high program FR and high vendor influence. The interviews will result in project specific vendor attribution scores that quantify the program's influence on the vendor's recommendations to the customer. Vendor attribution is an indirect program influence on the participant's decision to implement energy saving measures. Where program influence on the vendor's recommendations is greater than program influence on the participant directly, the vendor score will be used. We will complete IDIs with up to 80 of these vendors. The outline for Participating Vendor IDIs is provided in Appendix F .

[30] See Appendix D.

Mitigation Strategies for Data Collection Risks

Several risks to the data collection efforts have been identified in earlier discussions with the TEC, EAC and OEB. The timing of data collection in the year, identifying informed respondents and the handling of optimization, operations and maintenance projects are concerns that have been singled out.

Timing

DNV GL recognizes the limitations of the calendar in conducting survey research. Holidays increase the difficulty in reaching individuals. DNV GL will take efforts to conduct the majority of data collection outside of peak summer and winter holiday periods. Typical survey protocol dictates that contact with a survey respondent should be attempted 6-8 times before being considered 'exhausted'; DNV GL will adapt survey protocols to ensure that contact with an individual is not attempted more than 2x in a given calendar week and 3x in any two weeks to ensure that holidays do not influence response rates. DNV GL will also consider implementing a data collection hiatus during the last week of the year should the evaluation remain in the field at that time.

Informed Respondent

For data collection efforts involving non-program staff (e.g., participant surveys, participant interviews, participant follow-up interviews, participating vendor interviews), DNV GL will include a question battery designed to ensure that only informed respondents are participating. For participating customer respondents, DNV GL will define informed respondents as interviewees who directly participated in the project(s) in question. For participating customer respondents, DNV GL will make every effort to reach informed respondents.

To ensure informed respondents we will provide the programs the opportunity to verify that the program tracking data an appropriate contact listed. A spreadsheet listing the sampled projects (and backups) will be provided and the utilities will be asked to identify a "decision making" contact (for NTG) and a "technical contact" (for CPSV). In some cases these will be the same person. For TSERS where separate contacts are listed for the two roles, we will complete the separate parts of the IDI with the appropriate contact for that part.

When we call each contact we will list the project(s) that we are asking about, providing a description of the measures, location and time frame of implementation. We will then ask "Are you familiar with your organization's decision to make these energy efficiency improvements?" if the respondent indicates "yes" we will follow up to ensure that the "yes" response applies to all of the projects we are asking about.

For any projects that the response is "no" we will ask for contact information for someone who is familiar. We will continue with the survey for the projects that the respondent indicates that they are familiar with.

Some companies with multiple projects and diverse decision makers may require multiple interviews. We will not administer surveys for projects where the informed respondents are not available.

Response Rates

Survey response rates have been in decline over the past decade. This is especially true for residential surveys, where cell-phone only households have made surveying difficult, but there has also been erosion of response rates for business surveys. In order to achieve increased response rates, DNV GL will prompt program participants with both advance emails and advance letters, informing them of the survey and requesting participation. Advance letters, sent through traditional postal mail, are generally better received (and read) when sent by the recognized energy provider and should be sent on utility letterhead, if possible.

All communications with program participants will adhere to each of the respective utility's protocols for customer communication.

In order to execute the mailings, it is critical that DNV GL be provided with accurate contact information for the correct informed respondent. This will include, but is not limited to, the correct individual's:

- Full Name
- Role
- Mailing Address
- Email Address
- Direct Business Phone Number

DNV GL will send the above-mentioned emails and letters to all program participants included in the primary and backup samples of each data collection effort. For IDIs, there is an additional opportunity to improve response rates – providing respondents with the opportunity to schedule their own interview time. DNV GL will accomplish this with either an invitation to email DNV GL directly about preferred times or will utilize an online scheduling service where individuals may choose their own preferred times.

Secondary Attribution

Optimization, operational and maintenance projects (Actions) will be separated from equipment installation in the sample design and require special consideration for data collection as well. Maintenance projects in particular are by their nature recurring, while equipment optimization and operational improvements are behavioural and can be undone. The question of how to credit the program for maintenance this year when the customer participated in the past is complex. DNV GL and the TEC considered this issue while finalizing the contract and decided that the primary objective of the free ridership estimation will be to capture the effect of the program(s) on the current project. The effect on the current project of prior and indirect program experience will be captured in a secondary, less rigorous question sequence.

The primary attribution questions will be framed by questions that ask about decision making for the current project alone so that the scored attribution sequence will capture the effect of the program on the current project. After the scored section of the survey is complete we will capture the indirect, longer term attribution effect by asking:

- "Now, without *any* utility assistance for *any* projects in the past, what is the percent likelihood that you would have <taken this EE Action>?

The maximum of the primary attribution and this score will provide us with an idea of how much higher attribution would be if a longer term view were taken.

To limit customer burden and ensure the validity of our spillover analysis we will limit the investigation of secondary attribution to:

- *Measures with less than 100 percent primary attribution*: if primary attribution is 100 percent, then secondary attribution is as well. Put another way, the long term effect of the program (secondary attribution) is inclusive of short term (primary attribution), so by definition, the long term effect cannot be less than the short term and it is unnecessary to ask the secondary attribution question.

- *2015 participants*: 2015 participants are the only participants that will be asked attribution questions comprehensively. 2013/14 participants will be asked the attribution questions only if they indicate potential spillover has occurred (potential spillover is a project that the participant reports as possibly reducing gas usage, not-incentivized and attributable to previous program participation.

Secondary attribution will be captured for all measure types, but was a specific concern for maintenance and other “Action” measures.

Spillover Concerns

Key challenges to providing convincing quantified evidence of spillover for a particular customer include:

- Determining that a particular subsequent action was due to the influence of the program
- Confirming that the action was not taken as part of the original or another program, hence already counted by the program
- Quantifying the savings associated with confirmed spillover actions.

DNV GL’s approach provides a high level of rigor to address each of these issues.

- We determine program influence using participant surveys that start with the framing used for our free ridership questions. This framing helps ensure more meaningful responses to questions of the influence of the experience with the program in implementing the original measure on subsequent actions. As for the free rider surveys, obtaining the right respondents is also essential to obtaining meaningful responses to these questions.
- We confirm that the actions tentatively identified as spillover were not already counted by another program by cross-checking tracking data bases. Also critical to separation of spillover from program-claimed savings is understanding what savings if any are claimed by the programs for facilitation support such as opportunity identification, feasibility studies, audits, and related continuous improvement program engagement.
- We quantify the savings for confirmed spillover actions by collecting engineering specifications and calculating associated savings. This approach gives more accurate results than asking customers to estimate the magnitude of spillover savings relative to the original measure.

Thus, our participant spillover methodology addresses the following key issues:

- *Locating the right decision-maker* - Large commercial and industrial companies have multiple decision-makers and it is often difficult to find someone who is familiar with both the tracked program-influenced measure and the spillover measure. Employee turnover can also complicate this. Our approaches to ensure appropriate respondents are discussed above.
- *Avoiding double-counting* – Companies that received financial incentives from an energy efficiency program for one measure are likely to seek these incentives for future measures. Hence it is important to get the program’s latest tracking data to make sure that a potential participant spillover measure did not receive program support.
- *Estimating program attribution for potential spillover measures* - A common way of assessing participant spillover is to ask how much the participant’s experience with the tracked program-influenced measure influenced their decision to implement measures that are candidates for spillover attribution. It is difficult to turn this “fuzzy” assessment of program influence into a more concrete attribution factor necessary for attributing a certain quantity of m^3 from the spillover measure to the program.
- *Estimating the energy savings for the participant spillover measures.* Because spillover measures occurred outside the program, evaluators do not have access to the same information about the size, type, and quantity of the implemented energy-efficient measures that they would find in a program tracking database.

Task 4: Data Analysis

The data analysis task takes the data collected in Task 3 and combines it into adjustment factors that represent the population of implemented projects. Those adjustment factors are then applied to the program-level savings to produce verified gross savings and net savings. Table 26 shows the sub-tasks and their completion status for Task 4.

Table 26: Task 4 Subtasks and Completion Status

Key Tasks
☒ Task 1: Project Kickoff
☐ Task 2: Sample Design
☐ Task 3: Data Collection
☐ **Task 4: Data Analysis**
☐ **Task 4.1: Analyze Survey and Interview Data**
☐ Sample frame data transformation
☐ Sampling weight
☐ **Task 4.2 Calculate estimates**
☐ Verified Gross Savings
☐ Free Ridership (attribution)
☐ Secondary Attribution
☐ Spillover
☐ Gross 2015 program savings
☐ Net 2015 program savings
☐ Task 5: Reporting
☐ Task 6: Project Management

The objectives of this task are to:

- Determine the population-weighted adjustment factors related to verified gross savings, FR, spillover, and NTG
- Apply the adjustment factors to the appropriate program-reported savings estimates
- Produce the overall verified gross and net savings
- Produce the overall spillover adjustment factor

This task will begin with preliminary (incomplete) data collected in Task 3. The preliminary data will be used to establish the analysis methodology, which will be implemented once the data collection is complete. Each activity will be discussed in greater detail below.

Task 4.1: Analyze survey and interview data

The analysis flow after data collection begins with transforming the collected data back to the level of the unit of analysis. This translation depends on the number and grouping of program measures or projects asked about for an individual customer, and whether subsampling was required.

The survey will collect attribution information on each measure type. We apply the free ridership and spillover “scoring” methods to determine the free rider and spillover factors for each measure type. We then apply these factors to the associated gross savings to produce net-of-free riders and spillover savings for each measure type. Data collected from a single customer will be treated as a single cluster in error estimates.

We will use the sampling weights created during the sample design process to expand the customer sample in each sampling cell (stratum) to represent the full participant population in that cell. Targeted cells for which we are unable to obtain any responses will either be treated as not represented by the sample, or will be collapsed with other cells for sample expansion.

The application of attribution and spillover algorithms that convert survey and interview data into energy savings values ready for expansion involves consistency checks for each respondent. These checks utilize both questions directly used in the algorithms and verbatim responses that contain information on the reasoning of the respondent's responses.

Task 4.2: Calculate Estimates

The 2015 combined evaluation will result in verified gross savings and free ridership that are calculated for each evaluated unit of analysis and expanded to the population using the statistical technique of ratio estimation.

Verified savings will be estimated by evaluation engineers while free ridership estimates will be calculated using the survey data collected. Free ridership will be calculated using the LCNS method by scoring survey responses as described in this section.

The 2016 spillover estimates will be calculated by a combination of evaluation engineer estimation and scored survey responses. Spillover will be captured using an approach that will capture inside, outside, like and unlike spillover. Separate estimates for each of the four categories will be produced as a ratio of spillover to gross savings.

More detail on this task is provided in Appendix D and Appendix E .

Task 5: Reporting

The reporting task encompasses the formal communication between the DNV GL team and the OEB and other stakeholders. Reporting includes status and update reports as well as the draft and final reports, which take the results of the analysis from Task 4 and presents them to the OEB, EAC, and other interested stakeholders. Table 27 shows the sub-tasks and their completion status for Task 5.

Table 27: Task 5 Subtasks and Completion Status

Key Tasks
☒ Task 1: Project Kickoff
☐ Task 2: Sample Design
☐ Task 3: Data Collection
☐ Task 4: Data Analysis
☐ Task 5: Reporting
☐ **Task 5.1: Reporting**
☐ **Task 5.1: Monthly Status Reports**
☐ **Task 5.2: Bi-Monthly Updates**
☐ **Task 5.4: Draft Deliverables**
☐ 2015 CPSV and FR Report
☐ Participant Spillover Report
☐ 2015 Verified and Net Savings Memo
☐ **Task 5.5: Final reports and presentation**
☐ Final reports addressing comments on draft report
☐ In-person presentation
☐ Task 6: Project Management

The objectives of this task are to:

- Provide clear and timely reports on project progress and evaluation activities
- Distribute and document results
- Document methodology
- Make recommendations for program improvements

Task 5.1: Monthly Status Reports

Every month the DNV GL project manager will submit a status report to the OEB, via email, which will summarize the past month's activities, notify of the next month's activities, and report on how closely the evaluation is adhering to the original schedule. However, if there are methodological questions or delays in responses to data requests that could put the evaluation off schedule, the program manager will notify the OEB of these issues immediately for proposed resolution so that the evaluation schedule is not compromised.

Task 5.2: Bi-monthly Status Updates

The DNV GL project manager will provide the OEB with study updates via teleconference on a bi-monthly basis in alignment with scheduled EAC meetings. These bi-monthly study updates will provide similar information as in the monthly emailed status reports, although the more interactive format of the teleconference should allow for greater discussion and quicker resolution of any key issues.

Task 5.3: Draft Reports

At the conclusion of the evaluation, DNV GL will submit to the OEB two draft reports and one draft memo that will present all the information in the research objectives. Each of the draft reports and the memo will have separate results sections for each utility with common methodology sections. This will allow for streamlined review of sections that apply to both utilities, while facilitating a potential separation of each deliverable into utility-specific final deliverables.

The first report (2015 CPSV/FR Report) will include

- Verification rates by market sectors, programs and domains of interest with associated precision estimates for both the Enbridge and Union Gas' 2015 programs
- Free ridership factors by market sectors, programs and domains of interest with associated precision estimates for both the Enbridge and Union Gas' 2015 programs
- Along with these key findings, we will also show how these estimates were derived and what data from the IDIs and onsites were used to inform these estimates, including any qualitative findings regarding non-incentive based utility services.

The second report (Spillover Report) will include

- Estimates of participant inside and outside, like and unlike spillover[31] by market sectors, programs and domains of interest with associated precision estimates for both the Enbridge and Union Gas' programs
- Guidance on the development of a strategy for applying spillover data collected on previous program participation to forward looking DSM program activity
- Along with these key findings, we will also show how these estimates were derived and what data from the CATI survey and follow up IDIs were used to inform these estimates, including any qualitative findings regarding non-incentive based utility services.

The memo (2015 Verified and Net Savings Memo) will include

- Verified and net savings (including spillover) for Enbridge and Union Gas' 2015 Custom programs and RunitRight.

Task 5.4: Final Report and Presentation

After receiving comments on the draft reports from the OEB and EAC, DNV GL will produce final reports (possibly separate final reports for each utility, depending on filing requirements) which address all these comments along with a comment matrix that shows how we addressed them and why. We also plan to deliver an in-person presentation of the results to the OEB and EAC.

[31] Potential electric and/or water spillover savings will not be reported in kWh, but descriptively, as electric spillover is outside the specific scope of this evaluation.

Task 6: Project Management

The project management task is an ongoing task to ensure proper implementation of the project, including the schedule, budget, and scope. Table 28 shows the sub-tasks and their completion status for Task 6.

Table 28: Task 6 Subtasks and Completion Status

Key Tasks
☒ Task 1: Project Kickoff
☐ Task 2: Sample Design
☐ Task 3: Data Collection
☐ Task 4: Data Analysis
☐ Task 5: Reporting
☐ **Task 6: Project Management**
Task 6.1: Complete evaluation on time, on budget and within scope.
Task 6.2: Keep client informed

The objectives of this task are to:

- Ensure timely and on-budget deliverables
- Keep clients informed of project progress

This task is ongoing over the course of the project, and includes budget and workflow tracking, communication among DNV GL team members and partner firms, and invoicing. The subsequent sections discuss the project timeline and risks to effective project implementation.

Project Timeline

Our current schedule has the project completion as May 25, 2017. This schedule includes four extra weeks for data collection to accommodate the winter holiday season. EAC review periods are assumed to be 1-2 weeks depending on the specific deliverable. Utility delivery of data is assumed to require two weeks following request.[32]

[32] With the exception of project files and contact information spreadsheets, which are being delivered a bin per week for four weeks for the CPSV/NTG and Spillover CATI contact information being provided in early January.

Table 29: Schedule of Deliverables

Task	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
SOW										
Final Prep										
Sample Design										
Review Process										
Surveys/Interview Planning										
Development (4 Guides, 4 Methods)										
Review Process										
Training/Coordination										
Documentation Request										
Contact Information Request										
Data collection										
Spillover Advance Letters										
Spillover CATI										
Spillover Follow up IDI										
CPSV & FR Advance Letters										
Energy Advisor Interviews										
CPSV & FR Interviews										
Vendor Interviews										
On-site Recruiting										
On-site Visits										
Analysis										
CPSV Site Report Development										
CPSV Site Report Review										
CPSV & NTG Analysis										
Spillover Engineering Calculations										
Spillover Analysis										
Reporting										
2015 CPSV & FR Report						D	F			
Spillover Report									D	F
2015 Net Savings Memo									D	F
Project Management										

Risks and Contingencies

The tables in this section document the risks to project schedule, finances, and quality and the contingencies DNV GL has in place to handle them.

Table 30: Schedule Risks

Schedule Risks	Explanation	Contingency
Data Reception	Timing Controlled by Union/Enbridge.	1) The data required and will be requested once spillover CATI's are completed.
Documentation Reception	Timing Controlled by Union/Enbridge	1) Send formal documentation request with explicit, agreed upon deadline for documents needed.
Contact information Reception	Timing Controlled by Union/Enbridge	1) send worksheet for contact information request 2) include clear directions for worksheet completion, including context of what we are attempting to learn from the interviews. 3) ensure the worksheet is simple and easy to complete.
Contact information processing	Contact information may be incomplete or come in hard to use format	1) Clear directions for the request 2) use experienced analyst to prepare data for survey
Resourcing	Having the right resources available at the right time is a challenge with projects that have experienced delays	1) reserve necessary resources for project in DNV GL's internal systems. 2) keep project on schedule to avoid conflicts with other project needs. 3) keep project sponsor aware of needs and championing project
Survey House Availability	Availability at the right time is a challenge with projects that have experienced delays	1) Malatest has been contracted for the work
Review Periods	Dependent upon OEB/EAC priorities	1) establish clear and explicit deadlines for reviews
Decision Making	Dependent upon OEB/EAC priorities	1) schedule meetings with clear agendas that have key decisions up front.
Response Rates	Response rates on surveys have been declining, which can extend the time required for data collection	1) IDI rather than CATI for the CPSV and FR portions of the study 2) email participants prior to call to ask for cooperation 3) send advance letter to participants prior to call to ask for cooperation

Table 31: Financial Risks

Financial Risks	Explanation	Contingency
Currency Exchange Rates	USD/CAD rates have been highly variable	Fix prices in USD and/or CAD have adjustment process in the contract.
Resourcing	Planned resources have been promoted since project scoped and now cost more. Planned resources have left company	1) substitute with acceptable alternatives. Provide OEB/EAC with CVs once resources are finalized.
Timeline	Longer timeline tends to use project funds more than shorter timeline	1) seek to reduce schedule delays 2) Ensure efficiency or delay non-critical work when critical path is delayed to avoid additional expense

Table 32: Quality Risks

Quality Risks	Explanation	Contingency
Response Rates	Response rates on surveys have been declining, which can reduce sample sizes, introduce uncertainty about bias and make it hard to get data from large customers who have a large effect on final result	1) attempt a census so that call order does not matter. 2) IDI rather than CATI for the most complex and large projects 3) email participants prior to call to ask for cooperation 4) send advance letter to participants prior to call to ask for cooperation
Informed Respondents	Multiple people in a business are often involved in the decision to purchase capital equipment or spend money on optimizing or maintaining existing equipment. For consistency and cost reasons a single respondent from a company is preferable to interviewing multiple people at a business about the decision. Ensuring we have a respondent who knows enough about the decision to complete the project and the influence of the program on that decision is the crucial challenge of the data collection effort	1) Clear guidelines and screening questions to determine an informed respondent 2) removal from study of un-informed respondents 3) single interview for a project may require contacting multiple people at the site to determine an informed respondent.
Engineering Estimates	Spillover estimates will be based on engineer estimates of savings for projects that were not part of a program. We expect that these projects will not have the typical amount of documentation that we see for program incentivized projects. The engineering estimates will be based on respondent provided information, and in some cases may not include specific sizes or operating characteristics.	1) Engineers will be required to thoroughly document information collected from the respondents and from third party sources. Justification for savings estimates will be provided, along with values and sources of key assumptions and calculation methods. 2) A senior engineer (Tammy) will review all estimates. 3) transparently provide documentation of project savings (within confidentiality limits) in appendix of report

APPENDIX A SAMPLING PROCESS

This appendix provides detail on the

1. high level process used in sampling
2. exploration of tracking data
3. definition of the unit of analysis
4. stratification decisions
5. 2015 FR and CPSV sample design
6. 2013/14 Spillover Sample design
7. Sample and backup sample selection

High Level Process

A sample is a collection of data items such as those collected through surveys, metering or onsite observation. A sample design is required when a sample does not include the entire target population. Most sample designs are driven by cost constraints (including schedule constraints), desired precision or both. The sampling process described here ensures that all bases are covered, ensuring optimal precision around estimates of interest for the data collected. The process we followed was:

1. **Identify Goals, Methods and Constraints**: for sampling, the goals consist of identifying the primary and secondary estimates of interest: what quantitative results are most important. Defining the data collection methodology –the process used to gather the data for the analysis – and the estimation method – the approach used to calculate the primary estimate of interest – is critical for defining elements of the design. Cost and schedule constraints surrounding the data collection and analysis then determine an upper bound for the sample size.

   - **Goals**: For this study the primary estimate of interest is the NTG ratio for each program. The NTG ratio is the parameter that we are targeting for 90/10 precision for each program.

     As will be described later in the methodology memo, we calculate the NTG ratio as

     NTG = (1-FR)*(1+SO).

     Since spillover tends to be small, this formulation is mathematically very close to the simpler formula indicated in the recent Ontario evaluations

     NTG = 1-FR + SO.

     We prefer the multiplicative formula as a more complete expression of the relationship between free ridership and spillover.

     Previous work in Ontario indicates that free ridership is on the order of 10% to 60% across program segments, 50% overall on a savings-weighted basis. Spillover is on the order of 5%. Because spillover is generally small, the precision of the full NTG will in most cases be close to that of the net-of-free rider factor, even with a modest spillover sample size.

   - **Methods and Constraints:** We are using two data collection methods, each of which have different costs associated. Due to cost constraints we must limit our use of on-sites to those projects where it will make the most difference in the estimate. These will be deployed on the largest and most complex projects as identified based on tracking data descriptions. TSERs will be used to collect the balance of the data that we do not have the funds to collect with Onsites. For smaller and simpler projects where the decisions made are more straight forward, TSER verification provides accurate data at a reasonable cost.

2. **Define the unit of analysis**: The unit of analysis is the level at which final estimates will be made. Some studies have multiple units of analysis: process evaluation results may be based on respondent level estimates, while impact evaluation results may be based on measure or project level estimates. Sampling units do not need to be the same as the unit of analysis, but identifying both early is crucial.

   We are using the same definition for our sampling unit. Most customers have no more than three projects in a given year, and most projects are of only one or two measure types, so that we will be able to inquire about all of these in a single survey or interview of reasonable length.

   We plan to ask each sampled customer about attribution for all of the customer's measures. Only a handful of customers have more than three (unit of analysis level) measures in 2016, with a maximum of six.

   For customers with large numbers of projects and measures, we will ask about groups of measures or projects. The groupings will depend on details of the types of measures and savings magnitudes.

3. **Identify the target population:** The target population is the universe of items that inferences and estimates are desired for. In the initial scope of the NTG study, the primary target population was defined as future programs of the same type. Having future program years as the target population has two implications for the sample design. First, the applicable error associated with our estimates is the non-finite population corrected error (described in our discussion of sample size below) which requires larger sample sizes for a given precision. Second, analysis by sub-domains such as measure types within the programs becomes more important. The measure mix in programs changes from year to year and typically NTG varies more across measure types than within. For more accurate estimates of net savings for future program years, applying measure type NTG ratios will be preferred to program as a whole NTG ratios. At this time the question of prospective vs. retrospective application of NTG results is unresolved. The final sample design is expected to result in precision levels sufficient for either application of the results.

4. **Establish the Sample Frame:** The sample frame refers to the list or mechanism from which the sample is drawn. A perfect frame will match the target population exactly.

   Since the target populations of this study are the future programs, we will not have a perfect sample frame; however, if the program designs remain relatively stable, using past program participants as the sample frame will provide a good list from which to draw our sample.

5. **Determine sample size:** Sample size refers to the number of items that are selected from the sample frame in order to draw inferences and create estimates about the target population. In stratified designs, sample sizes are determined for each stratum.

   Critical to the sample size determination is the error ratio for each sampling cell with respect to the ratio to the estimated. The error ratio for ratio estimation is the equivalent of the coefficient of variation for estimation of a population mean. Our experience with conducting similar NTG studies of commercial-industrial customers is that the error ratio for the free rider rate is between 0.7 and 0.8 within reasonably defined sampling cells.

   In determining these sample sizes, the number of customers in the full population is also important for two reasons. First, if we are trying to estimate a parameter for a finite population, the sample size required is reduced by the Finite Population Correction or FPC. Second, we need to consider the number of completed surveys we can realistically complete given likely response rates.

Use of the FPC is appropriate when the parameter of interest represents a particular population. This situation applies when we are determining the free ridership factor or spillover rate for a particular program and time frame. When we determine these factors for all future theoretical projects, it is arguably more appropriate to treat the sample drawn from recent participants as coming from an essentially infinite population. Thus, for projection to future years we generally recommend against applying the FPC.

6. **Stratification:** Stratification is the partitioning of a target population. Stratification is discussed in depth in the sample design section in the body of the Scope of Work.

7. **Sample Selection:** Sample selection refers to the process of obtaining the sample of units from the sample frame. If all units on the sample frame are selected then the design is referred to as a census or certainty sample. Otherwise units may be selected either randomly or non-randomly, depending on the evaluation goals, constraints and amount of acceptable bias. The sample selection process is a critical feature of the sample design and has a direct impact on the expected precision and bias of estimates. The optimal sample selection process for a particular project can vary greatly.

8. **Unit and Item non-response** Unit and item nonresponse are potential sources of bias, depending on the nonresponse mechanism and the level of nonresponse encountered. Unit nonresponse refers to the absence of information from an entire sampled unit. Item nonresponse refers to the situation where some data are collected, but not all, from a sampled unit. The nonresponse mechanism refers to the process that is causing the nonresponse. If the probability of responding depends on the data items being sought then the nonresponse mechanism is said to be non-ignorable. Otherwise it is called an ignorable nonresponse mechanism. Nonresponse bias tends to be greater when the nonresponse mechanism is non-ignorable and as levels of item nonresponse increase.

   There are various ways to address nonresponse in a sample. For example, weight adjustments are often used to account for unit nonresponse and item imputation techniques are often used to account for item nonresponse.

   If nonresponse levels are low and the response mechanism is thought to be ignorable then one could ignore nonresponse and simply create estimates among the respondents.

   We recommend treating unit nonresponse as ignorable for this study since it does not depend on the data items being sought. Instead, it depends on the willingness of the decision maker at the participating business agreeing to respond to the survey.

   For item nonresponse in the scored portion of the surveys we recommend treating the nonresponse as non-ignorable if all three of the T, E, Q portions of the free ridership sequence contain non-response. Otherwise we plan to treat the item nonresponse as ignorable and will impute the average response for the missing item from among scored units of the same measure type and utility. The exception to this rule is when we find conflicting responses in our QC of the data collection that indicates the nonresponse is non-ignorable. For non-ignorable item nonresponse we will drop the unit from the analysis.

9. **Expansion** Sample expansion refers to the process of extrapolating results from a sample back to the target population of interest. Often times this is done using a sample weight. The weight is a numeric quantity associated with each responding unit and conceptually represents the amount of the target population the responding unit represents during the analysis. The sample weight is some function of the total number of units on the sample frame.

   The sample weight for our analysis will be built from the inverse probability of selection, incorporating additional adjustment factors to account for nonresponse and coverage errors. The sample weight will be utilized along with the "size" of the unit (energy savings) to expand results using ratio estimation, as described in the ratio estimation appendix of this work plan.

10. **Domains of interest:** Often times, estimates for an entire target population are of interest, but so are estimates for various subgroups. Subgroups may or may not overlap. Identifying the population domains of interest is another critically important design feature because it affects the decisions being made about other design features, such as the desired sample size, stratification variables and primary and secondary estimates of interest.

Explore the Tracking Data

We explored the tracking data provided by Union and Enbridge to determine data availability, the number and types of measures installed, and the size and quantity of projects. We explored the Union and Enbridge datasets separately.

Enbridge Custom Participant Data

The custom program participant data files provided by Enbridge included custom C&I energy efficiency projects claimed during the 2013-2015 program years and custom Low Income Multi-Family projects claimed in 2015 (Table 11). The records in the tracking data are per installed measure, so there may be multiple rows per project if more than one measure is installed, and there may also be more than one project per account. There are 124 accounts that appear in both the 2015 program year and the 2013/14 program years.

Table 33: Enbridge Custom C&I and Multi-Residential Program Participation Metrics by Year

Program Year	Accounts	Gas Savings (m^3)
2013	680	53,030,333
2014	573	46,195,015
2015	706	51,330,067

The Enbridge custom project tracking data includes measure level savings specific to a measure, site and date. As part of defining the unit of analysis, we used the tracking data variables *Market Type, load type name*, *end use*, and *technology* to categorize measures into measure types that would be meaningful for data collection and expansion, shown in Table 34.

Table 34: Enbridge Participation Metrics by Measure, 2013-2015

Measure Type	2013-2014 (SO)		2015 (CPSV/FR)	
	Accounts	Gas Savings (m^3)	Accounts	Gas Savings (m^3)
HVAC	636	32,807,840	380	19,105,965
Controls	337	17,821,495	175	13,868,059
Other Equipment	121	25,151,192	10	2,153,339
Operational Improvements	119	9,672,787	55	7,811,661
Heat Recovery	16	1,092,519	29	4,398,419
Steam and Hot Water	175	3,376,999	86	1,825,048
Process Heat	14	4,786,413	3	73,078
Building Shell	38	1,833,941	89	1,794,104
Greenhouse	10	2,682,162	3	300,394

Enbridge RunitRight Participation Data

For RunitRight, the program tracking data includes projects claimed in the 2014-2015 program years. These projects were all completed in 2013-2014; savings for a project in the program do not get claimed until after one year of site metering is complete.

Table 35: Enbridge RunitRight Program Participation Metrics by Year Claimed

Program Year	Accounts	Gas Savings (m^3)
2014	45	625,088
2015	28	542,442

The RunitRight program has only one measure type. It also has several projects with negative savings. Negative savings (increases in energy use) are possible results from retro-commissioning projects, sometimes due to calculation method (billing analysis based savings without weather, occupancy adjustment or production adjustment) or due to actual increases in energy use. Negative saving measures need to be handled carefully in ratio estimation: high FR on large negative savings projects can result in overall program FR <0, which is not a valid result.[33] Our recommended approach to the problem is to produce and apply ratios with separate domains for positive and negative savings projects.

Union Custom Participant Data

The participant data files provided by Union included energy efficiency projects claimed during the 2013-2015 program years. The records in the tracking data are per installed measure, so there may be multiple rows per project if more than one measure is installed, and there may also be more than one project per account. There are 67 accounts that appear in both the 2015 program year and the 2013/14 program years.

Table 36: Union Participation Metrics by Year

Program Year	Accounts	Gas Savings (m^3)
2013	352	369,438,742
2014	392	285,752,549
2015	462	201,620,726

33 Free ridership on negative savings results in more program savings, rather than less.

We used the *project type*, *equipment type*, and *project category* variables in the tracking data to categorize measures. Our first step was to distill the combined information from the three fields into measure types that would be meaningful for data collection and expansion, shown in Table 37. The largest measure types (by cumulative savings) were maintenance, steam and hot water, and optimization.

Table 37: Union Participation Metrics by Measure, 2013-2015

Measure Type	2013-2014		2015	
	Accounts	Gas Savings (m^3)	Accounts	Gas Savings (m^3)
Maintenance	222	255,847,232	79	37,181,863
Steam and Hot Water	161	119,657,223	91	39,229,635
Optimization	91	94,790,733	28	16,936,421
Ag and Greenhouse	149	64,895,560	73	31,875,980
Heat Recovery	86	38,174,741	52	19,797,904
Other Equipment	56	27,104,377	13	20,653,141
Controls	78	16,785,704	128	13,267,526
HVAC	48	14,885,291	49	8,829,742
Process Heat	25	13,242,538	10	4,536,172
Building Shell	152	5,599,318	68	3,597,883
New Construction	19	3,714,489	5	4,589,777
Cogeneration	4	494,085	1	1,124,682

Define the Unit of Analysis

Following data exploration, we defined the *unit of analysis*, which established the level at which data will be analyzed but not the level at which it will be collected, which is the *sampling unit*. We further discuss this distinction and how the sampling unit is defined in the Task 2.5 section.

The definition of the unit of analysis is one of the most important and least discussed aspects of DSM program evaluation. Consider the following four dimensions: end -use, measure type, equipment or “action,” and calculation approach. The program tracking databases include the first three dimensions and do not have an identifier for the fourth (though there may be a way to proxy it). Our example assumes that calculation type can be defined at a high level with reasonable accuracy based on existing database fields for the 2015 program year. Table 38 shows six measures performed at a site in a year through a program. Each of these categories could be considered a possible unit of analysis.

Table 38: Example dimensions used to define a unit of analysis

Measure ID	Enduse	Measure Type	Equipment or Action	Calc Type
M1	Process Heat	Boiler	Action	Complex
M2	Process Heat	Boiler	Equipment	Complex
M3	Space Heat	Boiler	Action	Simple
M4	Space Heat	Boiler	Equipment	Simple
M5	Space Heat	Furnace	Action	Simple
M6	Space Heat	Furnace	Equipment	Simple

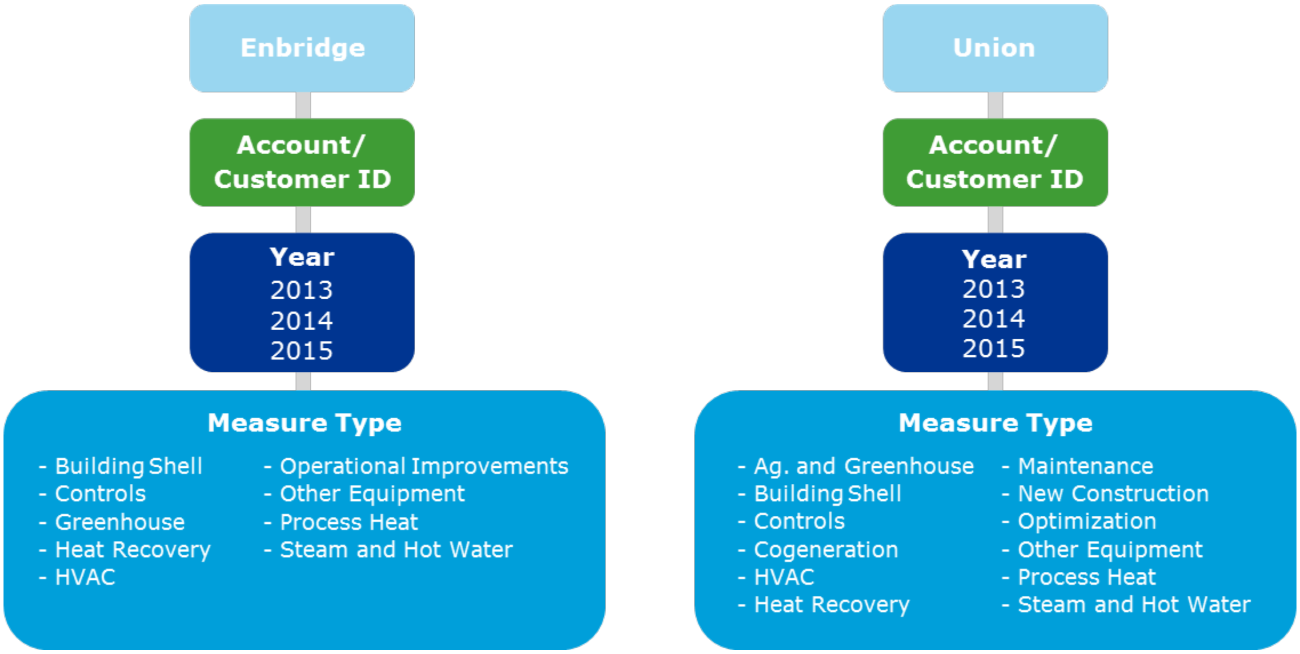
- The **end use** can be important in decision making because lowering the cost per unit produced is a different decision than lowering the cost of heating a facility or office, for example. It can also be used as a proxy for the complexity of the calculation, as process-related end uses tend to have more complex and site-specific calculation approaches. End use can be used in surveys by listing the measure types that fall into the category; however, this is not ideal for NTG as the program's influence on decision making may differ by measure type, affecting the attribution response.
- **Measure type** is important for surveys to aid participant recall by providing a concrete, simple description of what equipment was altered or installed. This aggregation is less appropriate for CPSV where the calculation method may differ.
- **Equipment or Action** is a very important distinction for NTG. Continuous improvement actions, such as maintenance, operations, and optimization, have fewer barriers to implementation than equipment purchases due to lower total cost, shorter term planning horizons and often fewer approvals. Businesses typically have separate budgets for capital and operating expenses. Purchases of new or replacement equipment falls under a capital budget, while actions are usually part of the operating budget or performed by salaried employees. Capital budgets typically have long term planning and allocation, while an operating budget is by nature more flexible to conditions in a given year. The ability of programs to affect equipment and action decision making is necessarily different as well. For the unit of analysis, actions were put into three categories: maintenance, operational improvement, and optimization.
- **Calculation type** is important for CPSV. Simple, commonly implemented measures in custom programs do not require the same depth of data collection to verify calculations and inputs as more complex measures. Simple measures also use standardized calculation approaches that reduce variance. Evaluators tend to find fewer adjustments and, even when adjustments are found, the adjustment often affects all measures of a calculation approach similarly.

In the example shown in Table 38, aggregating across any of the four listed dimensions is a trade-off of accuracy for increased precision, reduced customer burden and reduced evaluation costs. Not aggregating makes the same trade-off, but in reverse.

We aggregated across elements that are likely to have a lesser effect on decision making (such as type of insulation) and did not aggregate across distinctions that are likely to play a larger role in how decisions were made (such as process vs space heat).

The unit of analysis for the evaluation, presented in Figure 4, aggregates the data to the utility, account, year, and measure type. For Union, aggregating the tracking data to the unit of analysis reduced the number of records from 744 to 597 records for 2015 and from 1,468 to 1,091 records for 2013 to 2014. For Enbridge, the number of records for 2015 decreased from 955 to 858 records and for 2013 to 2014 decreased from 1,648 to 1,511 records.

Figure 13: Unit of Analysis



For this evaluation, the unit of analysis and the sampling units are defined differently. While a unit of analysis separates units of different accounts/sites, program years or measure types, the sampling unit is specific to the customer. As an example, one Enbridge customer may have installed a new boiler in 2013 and insulation in 2014, which is two different units of analysis. Since they were installed by the same customer, however, they belong to one sampling unit. In the analysis phase, weights will be developed for each unit of analysis (account-measure type-year), but for the standard error calculation, data collected from a single customer (sample unit) will be a treated as a cluster rather than evaluated as if they are independent observations.

Once aggregated to the unit of analysis, Union had an average of 1.5 units of analysis per account in 2013 and 2014 and 1.3 units per account in 2015[34] while Enbridge has an average of 1.2 units per account in 2013, 2014, and in 2015. In general, Union accounts tended to have more units of analysis per account than Enbridge accounts. Only 26 Union accounts have 5 units and none had more than 5. For Enbridge, 9 accounts have 4 units and no accounts have more than 4. This will facilitate data collection, since it's reasonable to ask about 3-4 units.

At this time we are unable to comment on the number of units per customer, because some customers will likely have multiple accounts. Customers will be defined by their contact information which will be requested along with the documentation request following submission of the scope of work.

[34] We are assuming a 1:1 account to customer ratio for sampling. For the analysis, customer will be defined by contact information (phone number primarily), which is not included in the provided tracking data.

Stratify the NTG and CPSV Data

There is a balance between having too many and too few strata.[35] In sample designs, more strata allow the design to control representativeness and estimated precision along more dimensions. Having more strata does not hurt overall precision, but it can increase the sample sizes required. Each stratification level serves to improve efficiency, improve representativeness, or both.

There are four populations across which the evaluation findings will be completely separate from one another.[36] These populations are defined by having separate program designs. The divisions between these populations are hard lines; none of the reported ratio results will include a mix of information across these populations. We can think of this as four evaluations using a common methodology and data collection effort.

5. Union Large Volume
6. Union Custom C&I
7. Enbridge Custom C&I
8. Enbridge RunitRight

Within the stratification segments (see Table 10) we categorize measures to improve the efficiency and representativeness of the sample. [37] The stratification for the 2015 data collection effort balances the needs of two studies, with the CPSV sample a subset of the NTG sample. Each has differing measure categorization priorities. [38]

- For NTG the measure categorization most predictive of free ridership rates is whether the project is installation of efficient equipment or whether the project was an action taken with existing equipment, regardless of whether that action is maintenance or an optimization that leads to energy savings.
- For CPSV the measure categorization most predictive of verification rates is a simple calculation versus one that is complex. Simple projects that follow consistent approaches and vary less from site to site typically have verification rates with lower variance than more complex projects that require more site specific knowledge and truly custom calculations. Stratifying by rigour allows us to assign a lower ER (0.3) to the simple project strata and higher to the more complex strata (0.4 ER) which provides better sample allocation. Simple strata projects will receive a TSER verification, while complex strata projects will receive an onsite verification.

The final stratification level segments projects by the magnitude of energy savings resulting from that project. Large projects represent a greater portion of the population, so sampling them at increased rates will result in greater precision with fewer verification visits or calls. Smaller projects must also be sampled to ensure representativeness. DNV GL used annual savings as a measure of size; cumulative savings were

---

[35] DNV GL agrees with the approaches described in "Sampling Methodology for Custom C&I Programs" which was prepared by Navigant for the TEC in 2012 and used to inform previous CPSV sample designs. Our sample design approach is consistent with the approaches recommended and follows the recommended seven step process (pages 17-23).
Dan Violette, Ph.D. & Brad Rogers, M.S., MBA, Navigant Consulting, Inc. "A Sampling Methodology for Custom C&I Programs," Prepared for: Sub-Committee of the Technical Evaluation Committee. November 12, 2012 (Revised October 28, 2014).

[36] For the CPSV, LI MF will be reported with MR MF either together with Custom C&I or as a separate Multi-Family domain, depending on final sample sizes and precisions.

[37] Page 14 in the Navigant report provides an explanation of the rationale for stratification.

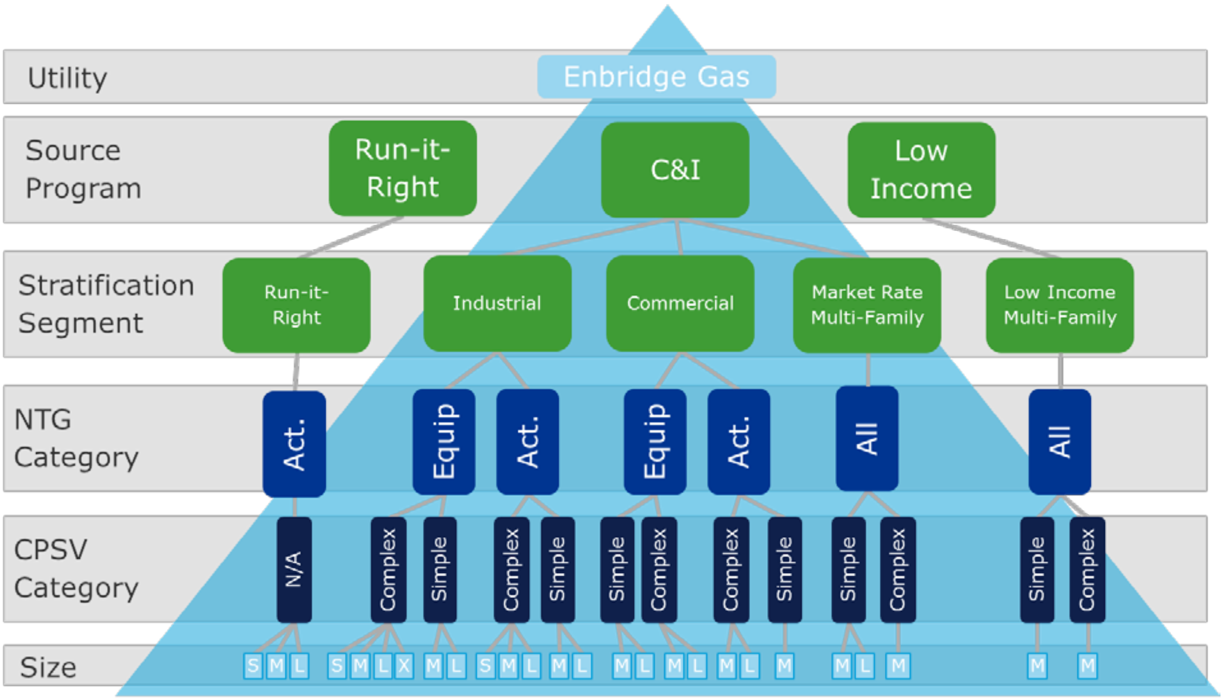
[38] The current stratification plan has more aggregated program segment categories than were described in the original proposal. When developing the proposal sample design we did not have access to the data or savings amounts specific to measure types.

not provided for all years and programs included in the study.[39] In terms of sample allocation, using annual savings will select longer life measures at a lower rate than would occur if cumulative savings are used.

It is important to note that the stratification used for sampling and expansion does not need to correspond directly to the level of reporting. For example, while we have chosen to use broad categories of customer segments in our stratification, this does not preclude reporting by more disaggregate customer segments.

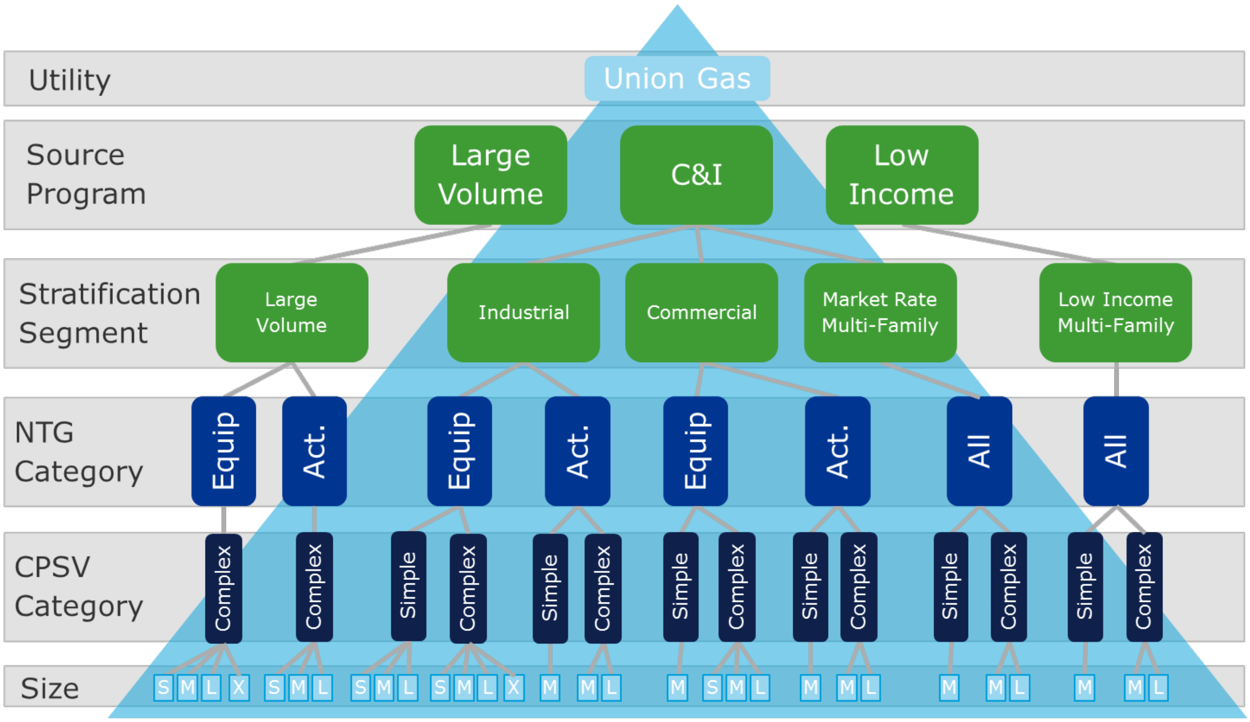
Figure 5 and Figure 6 show the CPSV and NTG stratification for Enbridge and Union respectively.

Figure 14: Enbridge Stratification



39 The August 4, 2016 data provided by Enbridge included cumulative savings for 2015, but not 2013 and 2014.

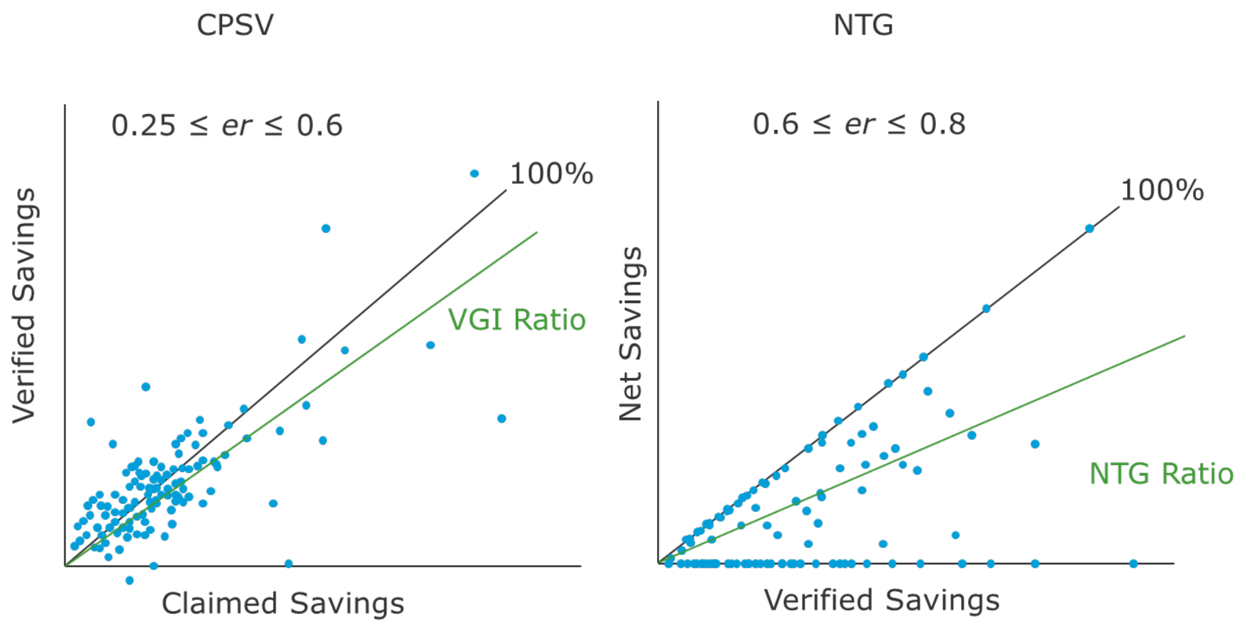
Figure 15: Union Stratification



Design the 2015 Samples

Critical to the sample size determination is the error ratio for each sampling cell with respect to the ratio to be estimated. The error ratio for ratio estimation is the equivalent of the coefficient of variation for estimation of a population mean. Free ridership is measured as a percentage between 0 and 100%, with clustering of responses on the extremes. The clustering of responses at 0 and 100% means that the error ratio for NTG studies is generally higher than that for engineering verification, where most of the estimates cluster reasonably close to the tracking savings estimates. Figure 13 shows the expected clustering of results for the two study types.

Figure 16: Error Ratio Example plots



Our experience with conducting similar studies of commercial-industrial customers is that the error ratio for NTG factors is between 0.6 and 0.8 within reasonably defined sampling cells. SO typically has an error ratio higher than that of FR. Our sample design assumes an ER of 0.6 for FR and 0.8 for SO.

Including the Equipment vs. Action level of stratification allows us to use a 0.6 ER assumption for FR, rather than the 0.7 ER assumption that we would use without.

The CPSV sample of the 2015 program year will target a subset of sites selected for the FR portion of the study. CPSV error ratios are typically lower than those for FR. We are using error ratios ranging from 0.3 to 0.4 for the CPSV portion of the study. Including a stratification level based on assumed complexity allows us to vary these ERs to better allocate our sample. Specifically, we used an error ratio of 0.4 for “complex” Commercial and Industrial strata, 0.35 for “complex” Multi-Family strata, and 0.3 for the less complex TSER strata.

The error ratios for CPSV are based on previous CPSV efforts for the utilities that have achieved or come close to achieving 90/10 precision at the program level using an error ratio assumption of 0.35. Using an error ratio from a study performed by a different firm working for different clients (even though they are the same programs) is a risk. We are mitigating this risk by using a unit of analysis smaller than site-level in our sample design, but collecting data on all projects at the site from the same program year. This approach provides an additional margin of error for the evaluation by collecting more data than is projected by the sample design approach.[40] The cost of the additional data collection is low since engineers will already be on the phone or onsite with the customer.

[40] Sampling at the sub-site level allows us to use measure characteristics more effectively in sampling and expansion. Over-collected data (units of analysis that were not selected randomly) will be given a weight of one (representing themselves alone) to ensure the final results are not biased by collecting additional data from multiple measure sites.

2015 Enbridge Stratification

The 2015 Enbridge stratification is presented in Figure 7. The final stratification includes 4 evaluation programs, two NTG categories, two CPSV categories and up to three size categories optimized for sampling efficiency. In total there are 26 Strata.

Table 14 shows the 2015 Enbridge sample design in table form with the expected relative precisions and sample for targeted programs shown for each portion of the study. Data collection for FR will be completed through IDI, while the CPSV group indicates strata that will have onsite data collection (complex) or TSER (simple) for gross savings verification.

Table 39: 2015 Enbridge CPSV and FR Sample Design

Stratification Segment	NTG Group	CPSV Group	Size Strata	Sample Frame		FR Sample Design		CPSV Sample Design	
				N	m3	n	Rel. Prec.	n	Rel. Prec.
Industrial	Action	Complex	3	18	3,839,353	8	9%	7	9%
		Simple	2	8	992,833	4		4	
	Equipment	Complex	4	66	15,781,190	24		14	
		Simple	2	26	2,193,173	5		5	
Commercial	Action	Complex	2	4	2,148,182	4	9%	4	9%
		Simple	1	24	824,845	4		4	
	Equipment	Complex	2	93	5,909,721	22		9	
		Simple	2	255	9,216,164	34		10	
Market Rate Multi-Family	All	Complex	1	55	2,139,551	8		5	
		Simple	2	182	5,224,012	19		8	
Low Income	All	Complex	1	6	297,979	0	N/A	2	
		Simple	1	93	2,763,064	0		6	
RunitRight	Optimization	IDI	3	28	542,442	18	10%	0	N/A
Total				**858**		**150**		**78**	

Table 15 shows the anticipated relative precisions for less aggregated program segments. We expect that the final relative precisions will be close to 90/10 for these segments as well as the targeted programs above.

Table 40: Enbridge Expected Precisions by Program Segment

Stratification Segment	Sample Frame		NTG		CPSV	
	N	Savings	n	Rel Prec.	n	Rel Prec.
Industrial	118	22,806,549	41	9%	30	9%
Commercial	376	18,098,912	64	10%	27	12%
MR MF + LI MF	336	10,424,606			21	13%
MR MF	237	7,363,563	27	20%		
RunitRight	28	542,442	18	10%		

Table 41 Provides the detailed sample design.

Table 41: Detailed 2015 Enbridge CPSV and FR Sample Design

Strata	Utility	Program	NTG Category	CPSV Category	Measures in Frame	FR Measure Target	CPSV Measure Target	Cumulative Gas Savings in Frame (m3)	Fraction of Frame Total Reported Cumulative Savings (m3)
211101	Enbridge	Industrial	Action	Complex	8	4	3	2,231,087	0.3%
211102	Enbridge	Industrial	Action	Complex	4	3	3	3,678,905	0.4%
211103	Enbridge	Industrial	Action	Complex	1	1	1	12,988,135	1.5%
211201	Enbridge	Industrial	Action	Simple	7	3	3	2,028,590	0.2%
211202	Enbridge	Industrial	Action	Simple	1	1	1	2,935,575	0.3%
212101	Enbridge	Industrial	Equipment	Complex	47	7	4	44,621,995	5.1%
212102	Enbridge	Industrial	Equipment	Complex	13	7	4	52,578,105	6.0%
212103	Enbridge	Industrial	Equipment	Complex	7	7	4	76,310,125	8.7%
212104	Enbridge	Industrial	Equipment	Complex	3	3	3	103,059,720	11.7%
212201	Enbridge	Industrial	Equipment	Simple	24	5	5	23,332,790	2.7%
212202	Enbridge	Industrial	Equipment	Simple	1	1	1	20,592,275	2.3%
221101	Enbridge	Commercial	Action	Complex	2	2	2	774,895	<0.1%
221102	Enbridge	Commercial	Action	Complex	1	1	1	10,213,885	1.2%
221201	Enbridge	Commercial	Action	Simple	24	4	4	3,875,430	0.4%
222101	Enbridge	Commercial	Equipment	Complex	50	13	4	20,106,586	2.3%
222102	Enbridge	Commercial	Equipment	Complex	8	8	4	31,966,255	3.6%
222103	Enbridge	Commercial	Equipment	Complex	1	1	1	9,501,060	1.1%
222201	Enbridge	Commercial	Equipment	Simple	265	17	5	88,190,023	10.1%
222202	Enbridge	Commercial	Equipment	Simple	28	17	5	148,466,935	16.9%
224101	Enbridge	Commercial	Multi-Residential	Complex	53	8	5	23,584,650	2.7%
224201	Enbridge	Commercial	Multi-Residential	Simple	139	10	4	53,999,911	6.2%
224202	Enbridge	Commercial	Multi-Residential	Simple	36	9	4	75,569,018	8.6%
241301	Enbridge	Run-it-right	Action	N/A	19	8	0	373,925	<0.1%
241302	Enbridge	Run-it-right	Action	N/A	5	5	0	923,845	0.1%
241303	Enbridge	Run-it-right	Action	N/A	4	4	0	1,414,440	0.2%
254101	Enbridge	Low Income	N/A	Complex	6	0	2	5,125,020	0.6%
254201	Enbridge	Low Income	N/A	Simple	104	0	6	58,676,555	6.7%

2015 Union Stratification

The Union stratification is shown in Figure 8. The final stratification includes 4 evaluation programs, two NTG categories, two CPSV categories and up to three size categories optimized for sampling efficiency. In total there are 30 strata.

Table 16 shows the 2015 Union sample design in table form with the expected relative precisions and sample for targeted programs shown for each portion of the study. Data collection for FR will all be completed through IDI, while the CPSV group indicates strata that will be have onsite data collection (complex) or TSER (simple) for gross savings verification.

Table 42: 2015 Union CPSV and FR Sample Design

Stratification Segment	NTG Group	CPSV Group	Size Strata	Sample Frame		FR Sample Design		CPSV Sample Design	
				N	m3	n	Rel. Prec.	n	Rel. Prec.
Industrial	Action	Complex	2	34	9,139,331	7	10%	7	9%
		Simple	1	26	4,882,678	4		3	
	Equipment	Complex	4	139	52,501,738	40		19	
		Simple	3	111	11,513,970	10		9	
Commercial	Action	Complex	2	5	6,776,934	5	10%	4	
		Simple	1	11	1,491,487	6		3	
	Equipment	Complex	3	113	6,850,571	17		8	
		Simple	1	42	1,013,521	5		4	
Market Rate Multi-Family	All	Complex	2	6	392,276	4		3	
		Simple	1	1	2,213	1		1	
Low Income Multi-Family	All	Complex	2	3	63,468	0	N/A	2	
		Simple	1	34	272,988	0		5	
Large Volume	Action	Complex	3	31	31,827,854	9	10%	8	10%
	Equipment	Complex	4	41	74,891,697	22		15	
Total				**597**		**130**		**91**	

Table 17 shows the anticipated relative precisions for less aggregated program segments. We expect that the final relative precisions will be close to 90/10 for these segments as well as the targeted programs above.

Table 43: 2015 Anticipated Precisions by Program Segment

Program Segment	Sample Frame		NTG		CPSV	
	N	Savings	n	Rel Prec.	n	Rel Prec.
Industrial	310	78,037,717	61	10%	38	10%
Commercial	171	16,132,513	33	10%	19	11%
Large Volume	72	106,719,551	31	10%	23	10%
MR MF+LI MF	44	730,945			11	13%
MR MF	7	394,489	5	18%		

Table 44 provides the detailed sample design.

Table 44: Detailed 2015 Union CPSV and FR Sample Design

Strata	Utility	Program	NTG Category	CPSV Category	Measures in Frame	FR Measure Target	CPSV Measure Target	Cumulative Gas Savings in Frame (m3)	Fraction of Frame Total Reported Cumulative Savings (m3)
111101	Union	Industrial	Action	Complex	21	7	6	75,487,148	2.8%
111201	Union	Industrial	Action	Simple	44	4	3	102,200,503	3.7%
112101	Union	Industrial	Equipment	Complex	104	13	6	183,932,142	6.7%
112102	Union	Industrial	Equipment	Complex	22	12	5	242,844,358	8.9%
112103	Union	Industrial	Equipment	Complex	9	9	5	347,468,949	12.7%
112104	Union	Industrial	Equipment	Complex	1	1	1	88,336,980	3.2%
112201	Union	Industrial	Equipment	Simple	91	5	4	50,638,424	1.9%
112202	Union	Industrial	Equipment	Simple	19	4	4	73,398,020	2.7%
112203	Union	Industrial	Equipment	Simple	1	1	1	41,029,840	1.5%
121101	Union	Commercial	Action	Complex	7	4	3	50,040,503	1.8%
121102	Union	Commercial	Action	Complex	1	1	1	31,595,400	1.2%
121201	Union	Commercial	Action	Simple	13	6	3	22,029,892	0.8%
122101	Union	Commercial	Equipment	Complex	104	9	3	20,998,185	0.8%
122102	Union	Commercial	Equipment	Complex	4	4	3	44,746,640	1.6%
122103	Union	Commercial	Equipment	Complex	1	1	1	76,886,900	2.8%
122201	Union	Commercial	Equipment	Simple	42	5	3	14,831,059	0.5%
123101	Union	Commercial	Multi-family	Complex	5	2	1	2,316,375	<0.1%
123102	Union	Commercial	Multi-family	Complex	1	1	1	5,093,140	0.2%
123201	Union	Commercial	Multi-family	Simple	1	1	1	44,260	<0.1%
131101	Union	Large Volume	Action	Complex	28	5	4	126,323,149	4.6%
131102	Union	Large Volume	Action	Complex	6	4	3	215,015,820	7.9%
131103	Union	Large Volume	Action	Complex	1	1	1	63,059,180	2.3%
132101	Union	Large Volume	Equipment	Complex	25	10	4	114,682,330	4.2%
132102	Union	Large Volume	Equipment	Complex	5	5	3	137,740,059	5.0%
132103	Union	Large Volume	Equipment	Complex	4	4	3	200,140,680	7.3%
132104	Union	Large Volume	Equipment	Complex	3	3	3	393,918,480	14.4%
153101	Union	Low Income	N/A	Complex	1	0	1	20,865	<0.1%
153102	Union	Low Income	N/A	Complex	1	0	1	1,433,430	<0.1%
153201	Union	Low Income	N/A	Simple	35	0	3	4,466,365	0.2%

Design the Spillover Samples

The sample design for spillover omits the CPSV category, but is otherwise consistent with the sample design for the 2015 FR and CPSV evaluation task. For spillover the ER used was 0.8; 90/10 precision was targeted.

2013/14 Enbridge Stratification

The 2013/14 Enbridge stratification is presented in Figure 7. The final stratification includes 4 evaluation programs, two NTG categories, two CPSV categories and up to three size categories optimized for sampling efficiency. In total there are 28 strata.

Figure 17: 2013/14 Enbridge Stratification

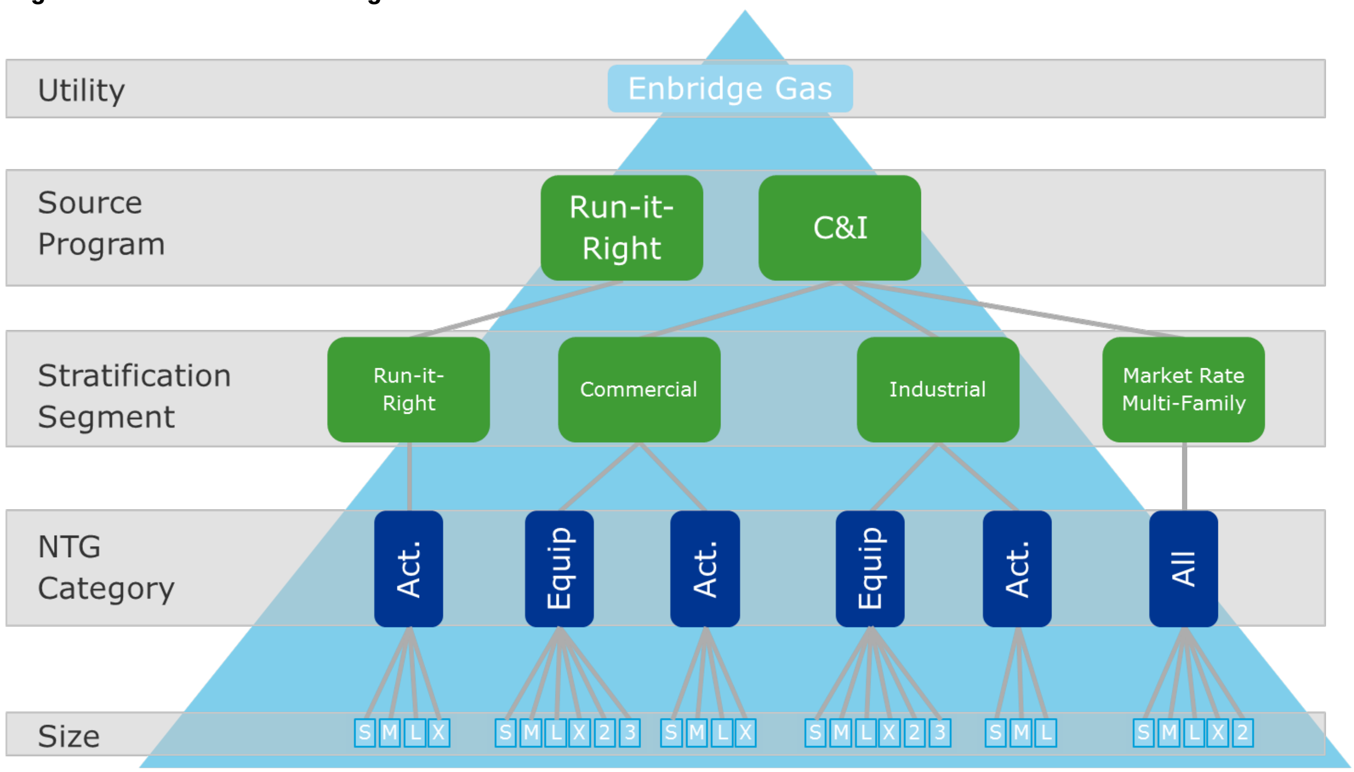


Table 18 shows the number of accounts and savings in the sample frame and the targeted spillover sample size for each grouping.

Table 45: 2013/14 Enbridge Spillover Sample Design

Stratification Segment	NTG Group	Size Strata	Sample Frame		SO Sample Design	
			N	m3	n	Rel. Prec.
Industrial	Action	3	40	5,067,923	20	10%
	Equipment	6	191	41,899,589	50	
Commercial	Action	4	79	4,604,864	25	10%
	Equipment	6	603	27,240,429	60	
MR MF	All	5	553	20,412,543	65	
RunitRight	Action	4	45	625,088	26	10%
Total			**1,511**		**246**	

2013/14 Union Stratification

The Union stratification is presented in Figure 8. In total there are 35 strata.

Figure 18: 2013/14 Union Stratification

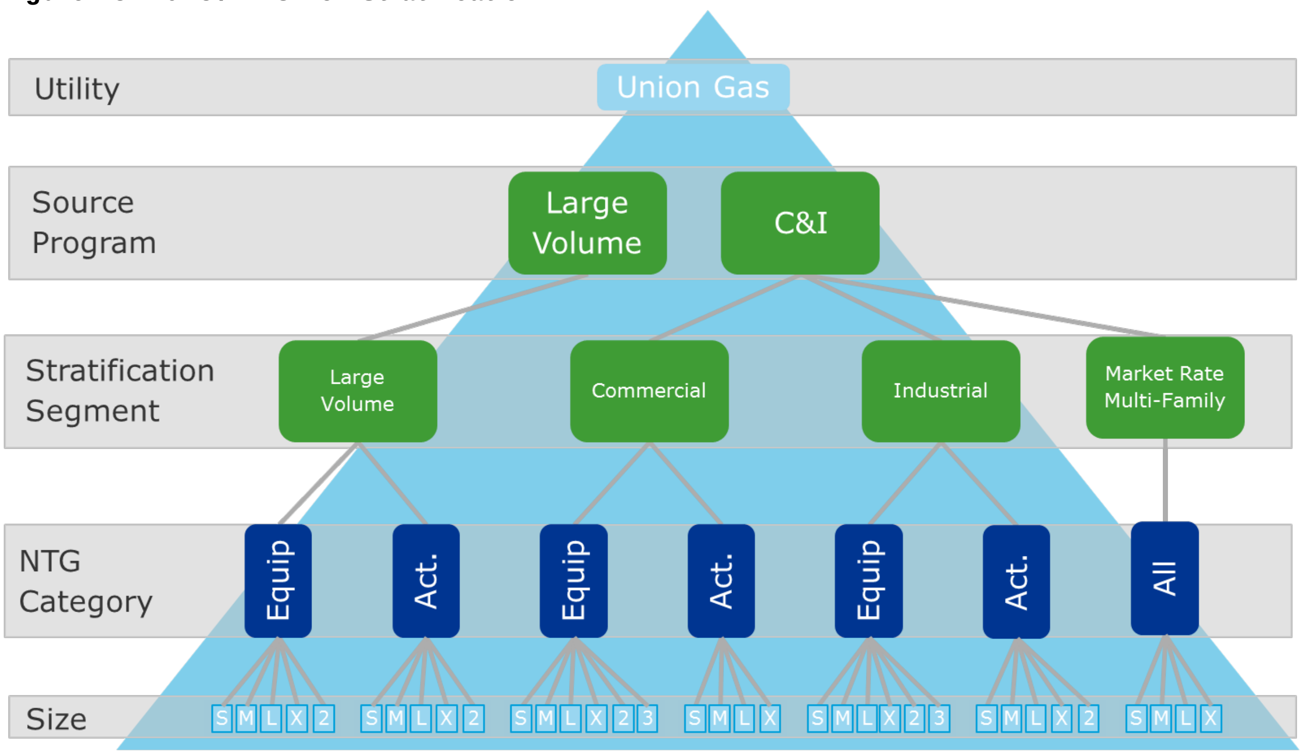


Table 19 shows the number of accounts and savings in the sample frame and the targeted spillover sample size for each grouping.

Table 46: Union Spillover Sample Design

Utility	Stratification Segment	NTG Group	Size Strata	Sample Frame		SO Sample Design	
				N	m3	n	Rel. Prec.
Union	Industrial	Action	5	167	64,448,800	38	10%
		Equipment	6	412	107,347,726	57	
	Commercial	Action	4	74	9,687,715	24	10%
		Equipment	5	190	15,744,760	40	
	MR MF	All	2	38	564,428	8	
	Large Volume	Action	5	130	317,638,812	38	10%
		Equipment	5	94	139,759,050	33	
	Total			**1,105**		**238**	

Prepare the Sample and Backup Sample

Once we have completed the final sample design, we will submit a data request to the utilities. The specific types of information we will be requesting are outlined in Table 47. The decision maker may not necessarily be located at the site where the project occurred and may be the same for multiple projects at multiple sites.

The technical expert is someone who will be able to answer questions regarding the specific engineering specifications of the equipment. Program energy advisors are the primary Account Manager or Energy Solutions Consultant that worked with the customer on the sampled projects. Vendors are the third party firms that were involved in the sale or design of the equipment, or the sale and performance of the O&M services.

Table 47: Information to Be Requested

Requested Information	Project Year	
	2013/14	2015
Site Address	√	√
Project Documentation	√	√
Decision Maker Contact Information: ▪ Full Name ▪ Role ▪ Mailing Address ▪ Email Address ▪ Direct Business Phone Number	√	√
Technical Expert Contact Information: ▪ Full Name ▪ Role ▪ Mailing Address ▪ Email Address ▪ Direct Business Phone Number	√	√
Program Energy Advisor Information: ▪ Full Name ▪ Email Address ▪ Direct Business Phone Number		√
Vendor Contact Information: ▪ Full Name ▪ Role ▪ Mailing Address ▪ Email Address ▪ Direct Business Phone Number	√	√

For the 2015 NTG sample we will request documentation and contact information for 50 percent more projects that are in the primary sample. This corresponds to a minimum 66 percent response rate. If response rates are lower than 66 percent in specific stratum, we will request documentation and contact information for additional projects in the stratum.

For the 2013/14 spillover sample we will request contact information for three times the number of sampled projects. This corresponds to a minimum 33 percent response rate. We will not request project documentation for the spillover sample until we have identified the sites that require follow up engineering interviews. To protect respondent confidentiality we will request documentation for more sites than will receive follow up calls. Overall this staging of requests will reduce the amount of project documentation that the utilities need to provide, while ensuring efficient data collection.

Backup sample will only be contacted if needed to meet targeted number of completes.

Once we have received the requested contact information, we will identify instances where a contact was involved in multiple projects, even across sites. While the projects are conducted at the site level, the decision maker, technical expert, or vendor may have been involved in projects at multiple sites. For example, multiple participating sites for the same retail chain may have one energy manager from the corporate office but the technical expert may be site specific. Using this contact information and taking into account cross-site involvement, we will assemble the CATI and the IDI sample frame.

APPENDIX B LCNS METHODOLOGY

Life Cycle Net Savings (LCNS) is a methodology for determining the FR component of NTG by estimating program effect over the life of the program measure. In this appendix, the terms FR and attribution are used interchangeably as complements of one another. This appendix does not include spillover.

Notation:

VGS_S = Verified Gross Savings based on standard efficiency equipment baseline (annual)

VGS_E = Verified Gross Savings based on pre-existing equipment baseline (annual)

VGS_L = Verified Gross Lifetime Savings

Y_L = Verified Estimated Useful Life (Years) of installed efficient equipment

Y_A = Years Accelerated

Y_R = Remaining Useful Life of pre-existing equipment

A_E = Efficiency Attribution

A_Q = Quantity (size) Attribution

F_E = Efficiency free ridership

F_Q = Quantity (size) free ridership

SPA = Simple Program Attribution (function of efficiency and quantity free ridership, not timing)

NS_L = Net Lifetime Savings

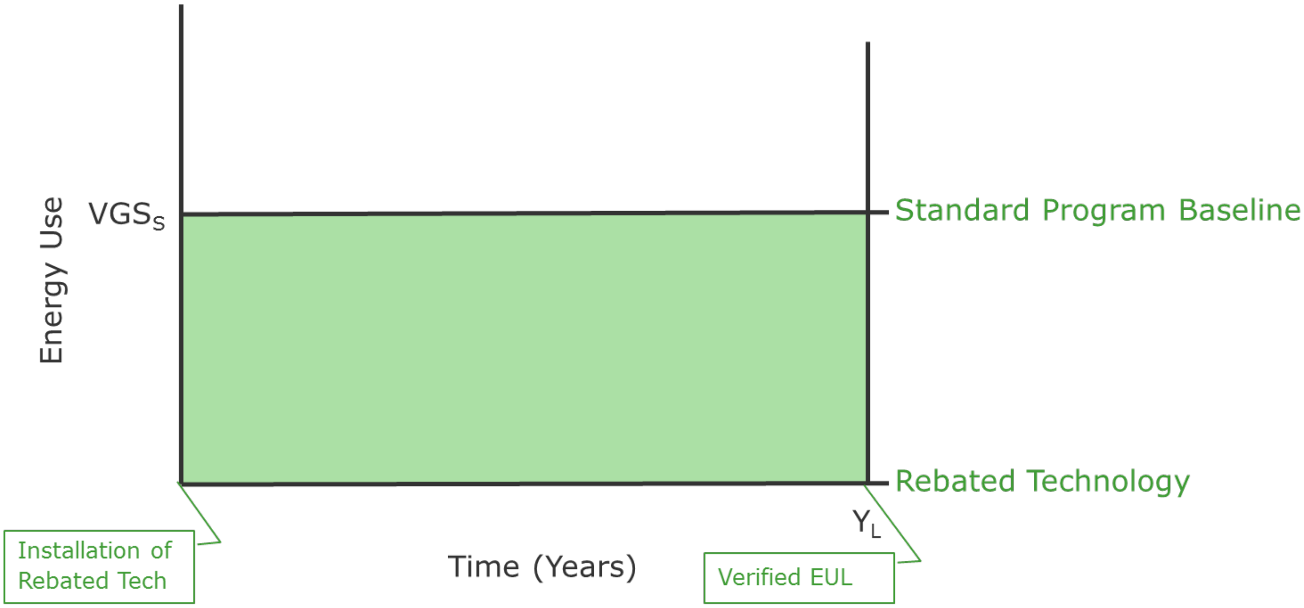
NS_A = Net Acceleration Period Savings

NS_P = Net Post-Acceleration Period Savings

Verified Lifetime Savings

First we consider the verified savings that make up the denominator in the NTG ratio. Figure 19 shows the verified lifetime savings for a measure.

Figure 19: Verified Lifetime Savings for a Measure



Verified lifetime savings are calculated as the difference in energy use of the incentivized measure and the energy use of a standard program baseline measure for the (verified) life of the measure.

$$VGS_L = VGS_A \times Y_L$$

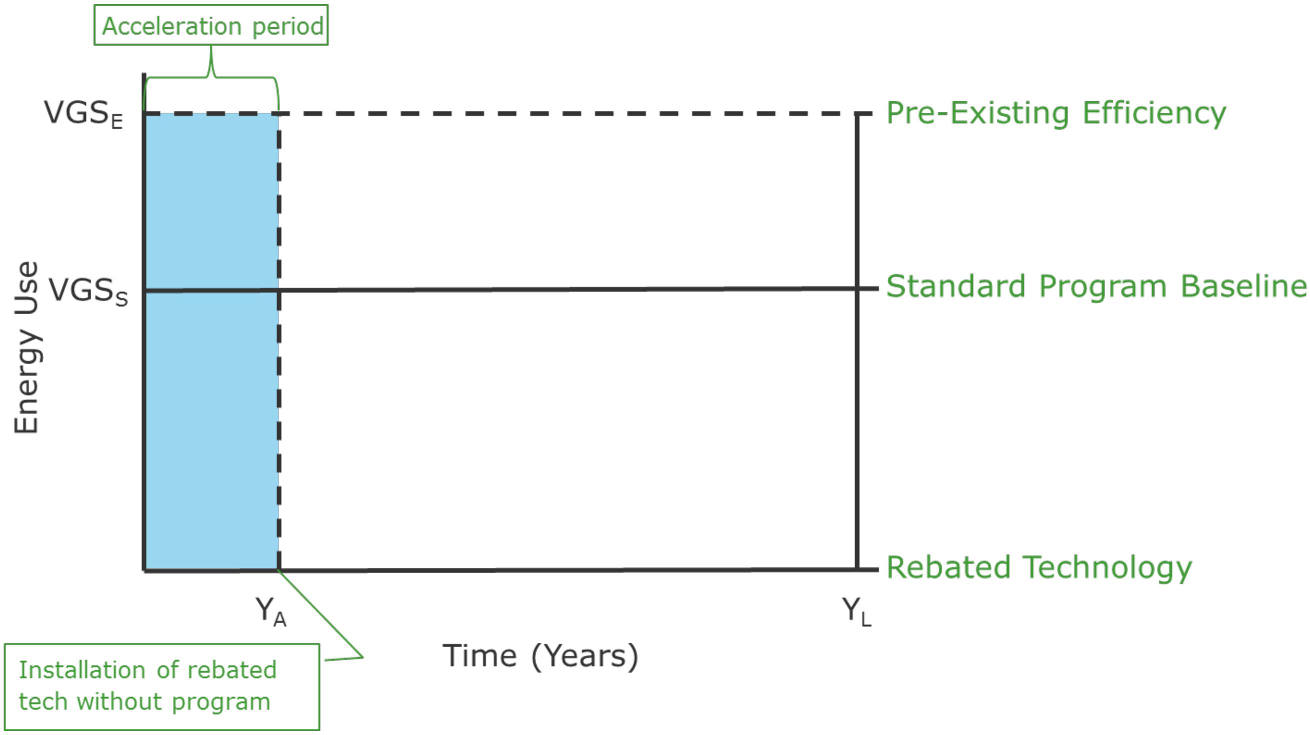
Timing

The treatment of timing is how LCNS differs from other estimation approaches for attribution. In LCNS the response to the question “when would you have performed the measure without the program” defines the number of years that the program accelerated (advanced) the measure. This period is referred to as the “acceleration period” and shown as the distance from the origin to Y_A along the x-axis.

During the acceleration period, the customer would not have installed a new measure (efficient or standard). Instead the appropriate baseline equipment for this time period is the pre-existing equipment that they had been using. This section shows how this difference in baseline affects the net savings estimate for the measure relative to the gross savings.

During the acceleration period (Y_A), the attributable savings are calculated as the difference in energy use of the incentivized equipment and the energy use of the replaced equipment (a pre-existing efficiency baseline). As a result, during the acceleration period the net savings (blue box up to VGS_E) may be higher than the verified gross savings (VGS_s) if the efficiency of the pre-existing equipment was less than the standard program baseline. Savings during the acceleration period are, by definition, attributable. Figure 20 shows the attributable savings in the acceleration period for an accelerated measure.

Figure 20: Acceleration Period Savings



Acceleration period savings are calculated as:

$$NS_A = VGS_E \times Y_A$$

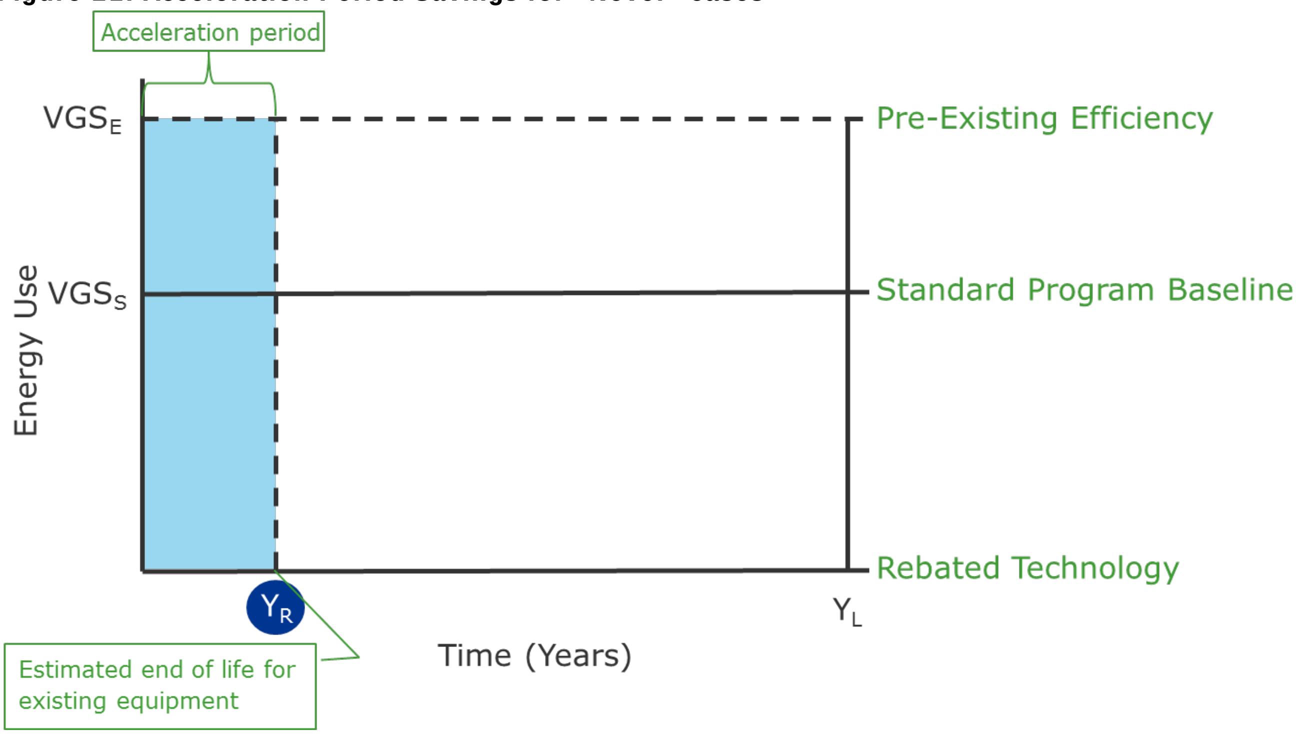
Special Case: "Never"

Some respondents will indicate that they would "never" have replaced the existing equipment. A customer "Never" would have installed the project if they:

1. respond to initial timing question by saying they never would have installed it without the program
2. respond to second timing question by saying they would have installed it more than four years later without the program
3. respond to the initial quantity question by saying they would not have replaced any of the units without the program

For these measures, the acceleration period is defined by the remaining useful life of the pre-existing measure (Y_R) and the applicable baseline is versus pre-existing efficiency (VGS_E) as shown in Figure 21.

Figure 21: Acceleration Period Savings for "Never" cases



Acceleration period savings for "Never" would have installed measures are calculated as:

$$NS_A = VGS_E \times Y_R$$

Efficiency and Quantity

In the post-acceleration period attribution is based on the program effect on the efficiency and quantity of what was installed.

Efficiency attribution, A_E, measures the effect the program had on the *efficiency* of the equipment installed. The efficiency attribution measures the proportion of savings attributable to the program for increasing the efficiency of the equipment above what would have been installed otherwise.

Quantity attribution, A_Q, measures the effect the program had on the *size or amount* of the equipment installed. The quantity attribution measures the proportion of savings attributable to the program for increasing or decreasing the quantity of equipment above or below what would have been installed otherwise.

The Simple Program Attribution (SPA) is the fraction of annual verified gross savings that are attributable to the program and is a function of the efficiency free-ridership (f_E) and the quantity free-ridership (f_Q).

The free-ridership values for efficiency and quantity are calculated from the attribution factors. The complement of attribution is free-ridership. Attribution measures the portion of the savings that result because of the actions of the program. Free-ridership measures the portion of the savings that would have happened in the absence of the program. The free-ridership equivalents of the attribution factors are used to determine program net savings.

$$f_E = 1 - A_E$$

$$f_Q = 1 - A_Q$$

The fraction of verified gross savings that would have occurred without the program is the product of the fraction of units that would have been installed without the program, and the fractional unit savings that these units would have had without the program.

$$f_{QE} = f_Q f_E$$

For example, if two-thirds as many units would have been installed without the program (f_Q = 2/3), and the savings per unit would have been only half as much (f_E = 1/2), the portion of the savings that would have occurred without the program would be

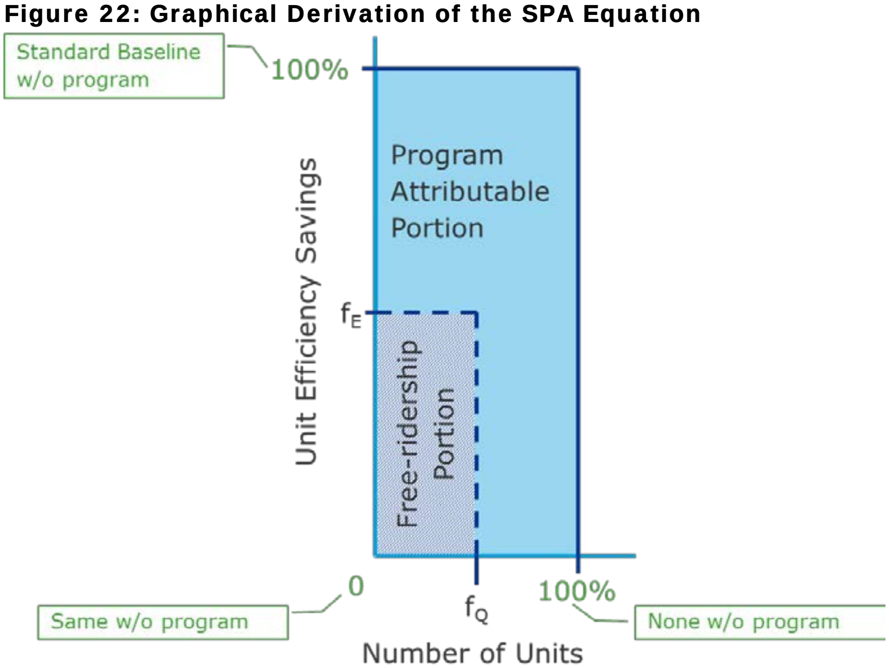
$$f_{QE} = (2/3) \times (1/2) = 1/3.$$

The SPA is the complement of this free rider portion.

$$SPA = 1 - f_{QE} = 1 - f_Q f_E$$

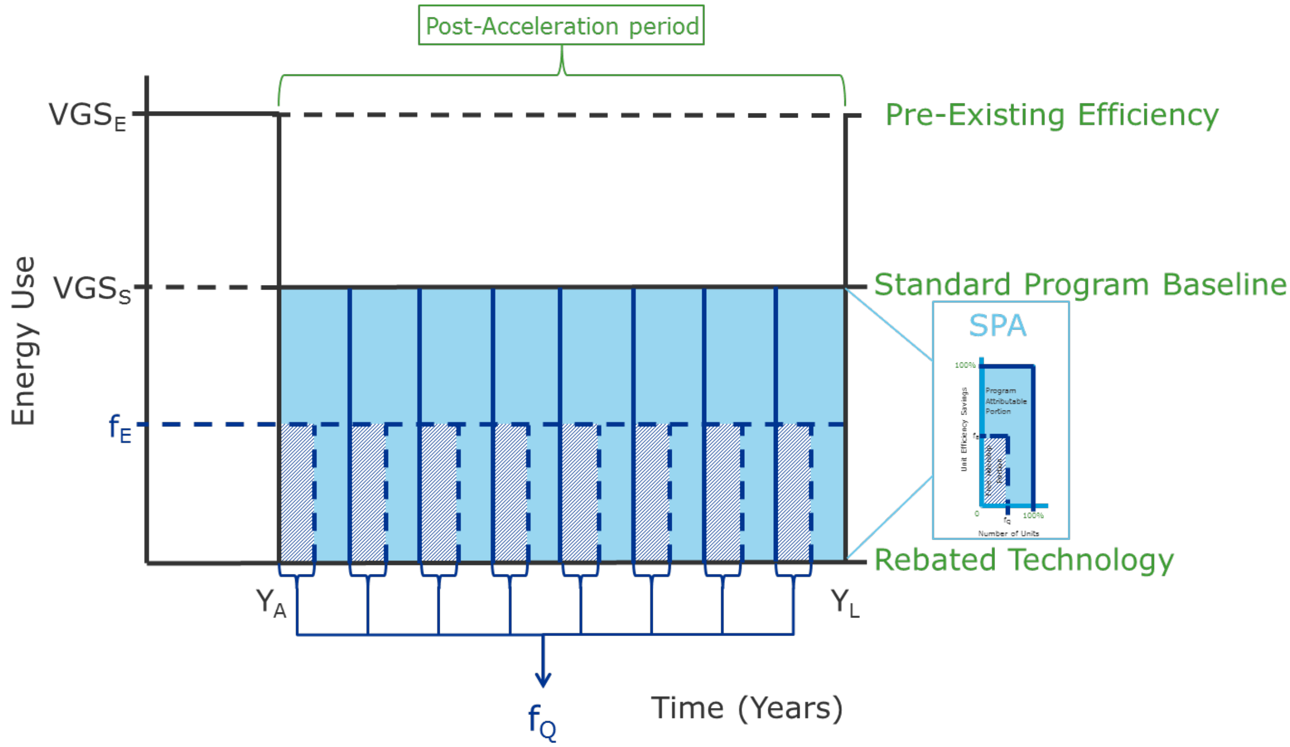
The relationship is illustrated in Figure 22.

Figure 22: Graphical Derivation of the SPA Equation



SPA is the attribution of each year savings in the post-acceleration period. Figure 23 shows the program attributable and free-ridership portions of each year's savings in the post-acceleration period. The blue rectangles represent SPA as discussed and shown from above. The height of the SPA box is equivalent to the baseline used for verified savings. The grey "missing pieces" are the free ridership for each year's savings. Because attribution is three dimensional and this is a two dimensional document, we are representing both years and quantity on the x-axis. Years are denoted by the dark blue vertical lines, while the quantity FR (f_Q) is shown as the width of the grey box.

Figure 23: Post-Acceleration Period Attributable Savings



The net savings in the post-acceleration period are calculated as:

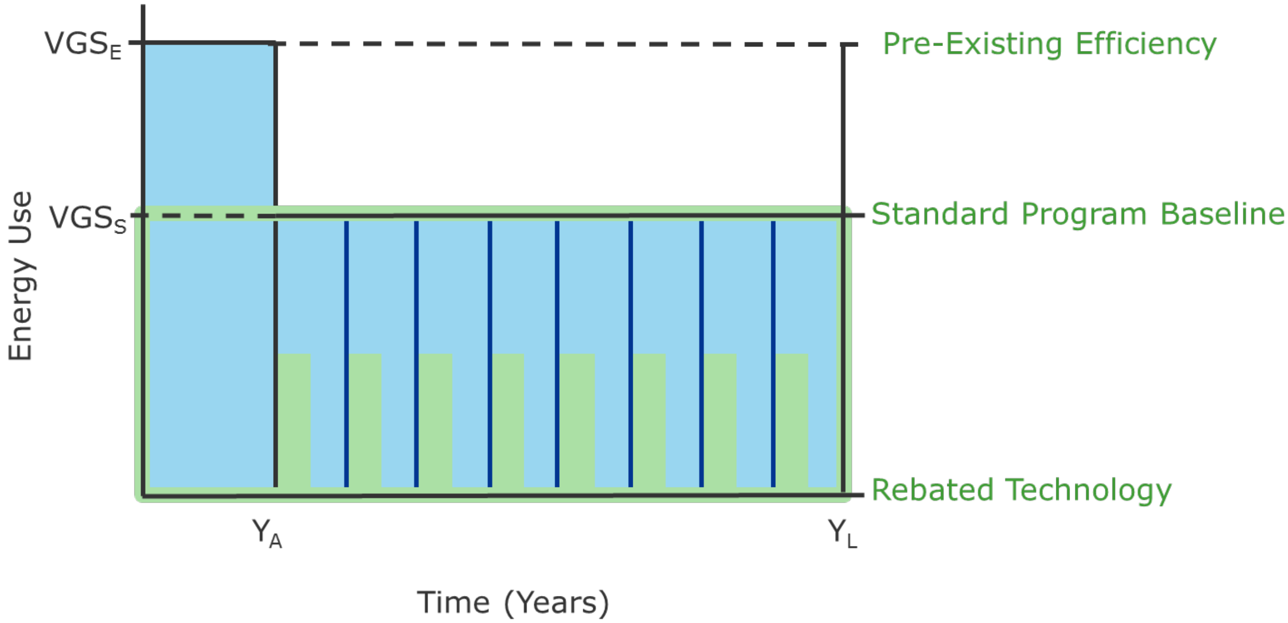
$$NS_P = VGS_S \times SPA \times (Y_L - Y_A)$$

Note that for the special case discussed relating to acceleration period savings, “Never”, SPA= 100%.

Calculating Attribution

Figure 24 shows the attributable savings across the lifetime of the measure NS_L (blue) overlaid on the verified gross lifetime savings VGS_L (green). The figure shows that with the effect of the dual baseline verification included in the net savings estimate and not in the verified savings estimate, some accelerated measures may have greater attributable savings than verified gross savings.

Figure 24: Attributable vs. Verified Gross Savings for a Measure



The formula for each individual measure's estimate of lifetime net savings is:

$$NS_L = NS_A + NS_P$$

or

$$NS_L = VGS_E \times Y_A + VGS_S \times (SPA) \times (Y_L - Y_A)$$

The formula for each individual measure's attribution is:

$$Attr = \frac{NS_L}{VGS_L}$$

$$Attr = \frac{VGS_E \times Y_A + VGS_S \times (SPA) \times (Y_L - Y_A)}{VGS_S \times Y_L}$$

Four years is the time horizon beyond which we assume the respondent cannot answer with certainty. Anything beyond four years (Y_A>=4) is treated as a "never would have installed" response (100% attributable), rather than an accelerated measure.

Special Case: FR Sampled Projects not sampled for CPSV

The sample for the CPSV portion of the study is a subset of the free ridership sample. This means that for projects included in the FR study, but not included in CPSV we will not be calculating verified savings. For expansion of the NTG ratio and for calculating post-acceleration period savings we will use the final ratio application domain level Gross RR to estimate verified savings for measures not in the CPSV.

For acceleration period savings we will use the A/P ratio of accelerated projects in the CPSV to estimate the pre-existing baseline savings. The A/P ratio refers to the ratio between the annual Acceleration Period Savings and the annual Post-Acceleration Period Net Savings. It is always one or larger. Like the application of Gross RR the A/P ratio will be estimated at the application domain level for use in estimating net savings for the FR-only sampled measures.

APPENDIX C DETERMINING ATTRIBUTION PARAMETERS

The attribution factors defined in the previous section are determined from the participant responses gathered during the survey. This section provides an overview of the survey data and how it is used to determine each attribution factor. It also includes more detailed sections for each factor that show exactly how all survey responses are handled.

General procedure

This section provides an overview of the attribution factors and how they are determined.

- Timing attribution, A_T: The timing attribution is determined from the acceleration period, Y_A, which is in turn provided directly by the respondent and the verified savings versus existing equipment provided by the evaluation engineers. There is no timing attribution effect for values of Y_A greater than four; in those instances we assume that the measure would never have been installed without the influence of the program.
- Efficiency attribution, A_E: The efficiency attribution is based on the answers to questions DAT2a and DAT2b which ask about the efficiency level that would have been installed in absence of the program. Respondents who indicate that they would have installed a lesser-efficient piece of equipment in the absence of the program are asked what efficiency they would have installed instead. An efficiency attribution value is assigned based on the response. Standard efficiency based on program definitions will be used to bracket the finer cut as defined in the project documentation provided by the utilities.
- Quantity attribution, A_Q: The quantity attribution is based on the percentage change in quantity caused by the program, ΔQ, which is in turn provided directly by the respondent. **Error! Reference source not found.** shows the attribution assignment based on responses to DAT3a and DAT3b.

The next few sections deal with determining the timing, efficiency, and quantity attributions on a more detailed level.

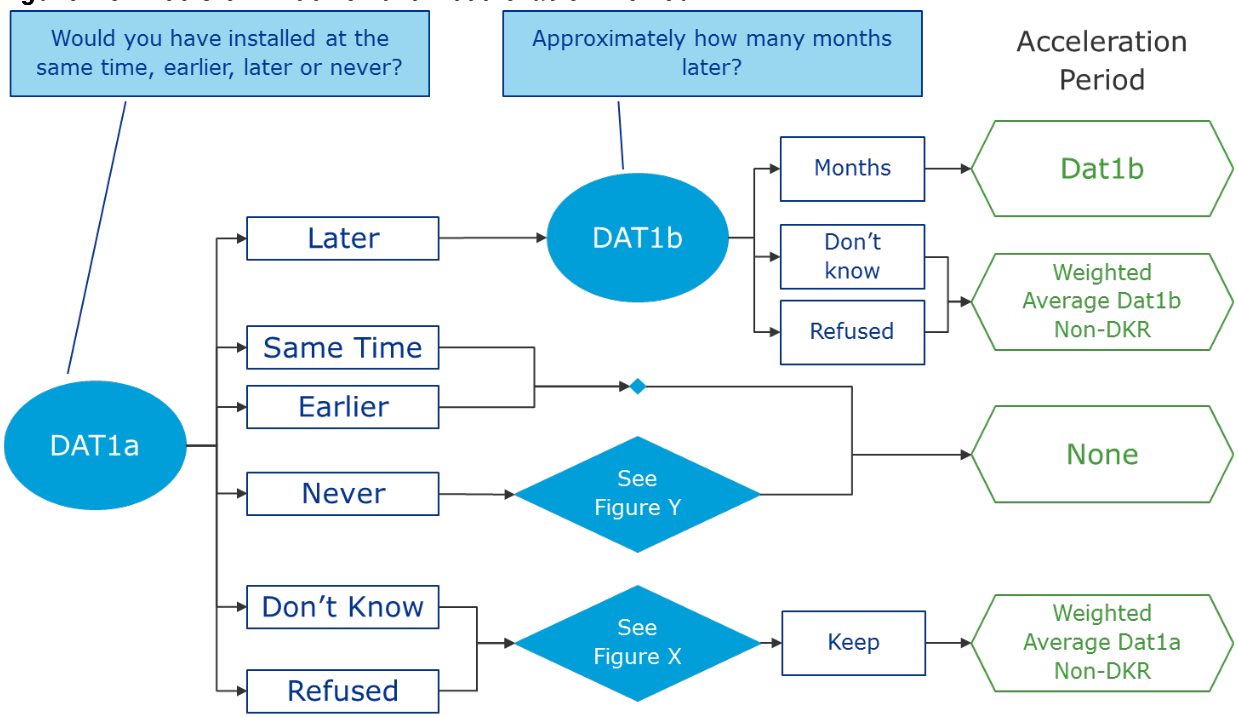
Timing

The timing attribution, A_T, is determined from the first set of attribution survey questions. These questions are used to determine whether or not the program accelerated implementation of a measure or caused it to be implemented before it would have been without the program. The two relevant questions are labelled DAT1a and DAT1b.

- DAT1a: "Without < the program>, would you have <installed, preformed> <measure> at the same time, earlier, later, or never?"
    - DAT1a_O: "Why do you say that?"
- DAT1b: "Approximately how many months later?" (DAT1b is only asked if DAT1a is "Later.")

Note that these questions ask about the timing of installing equipment, not installation of efficient equipment in particular. For example, if the measure was replacement of a high-efficiency boiler, the question asks when the boiler would have been replaced without the program. Engineers conducting the interviews are trained to ensure clarity for these questions. **Error! Reference source not found.** shows a decision tree for DAT1a and DAT1b.

Figure 25: Decision Tree for the Acceleration Period



The measure is considered accelerated if the respondent indicates that the measure would have been installed less than four years later without the influence of the program. The acceleration period is determined based on the answer to DAT1b. If the respondent is unable to answer DAT1b, the measure is assigned the average acceleration period across all accelerated measures in the same measure group.

If the respondent answers DAT1a with Earlier or Same Time then there is no acceleration period. If the respondent answers DAT1a with Never and the Quantity and Efficiency sections apply to the measure then the survey skips to the next section and there is no acceleration period. If the respondent answers DAT1a with Don't Know or Refused but does provide answers to inform the Quantity and Efficiency Attributions then the measure is assigned the average Acceleration Attribution for all measures in the same primary domain.[41]

[41] The primary domain is the domain that the attribution factor will be applied to in calculating the final net savings for the programs.

Table 48: Timing Attribution Assignments

Coarse Cut (DAT1a) (Would you have implemented the measure at the same time absent the program)	Finer Cut (DAT1b)	Acceleration period
Same time	NA	None
Earlier	NA	None
Later	0 < years <4	A_T=DAT1b Acceleration period equals response to DAT1b
	4<= years	Equivalent to “Never” $A_T=A_R$ Acceleration period equals remaining useful life of replaced equipment, SPA=100%
	Don't know/refused	Weighted average of "later" cases for primary domain, 0 < years <4
Never	NA	$A_T=A_R$ Acceleration period equals remaining useful life of replaced equipment, SPA=100%
Don't know/refused	NA	Weighted average of all respondents for primary domain

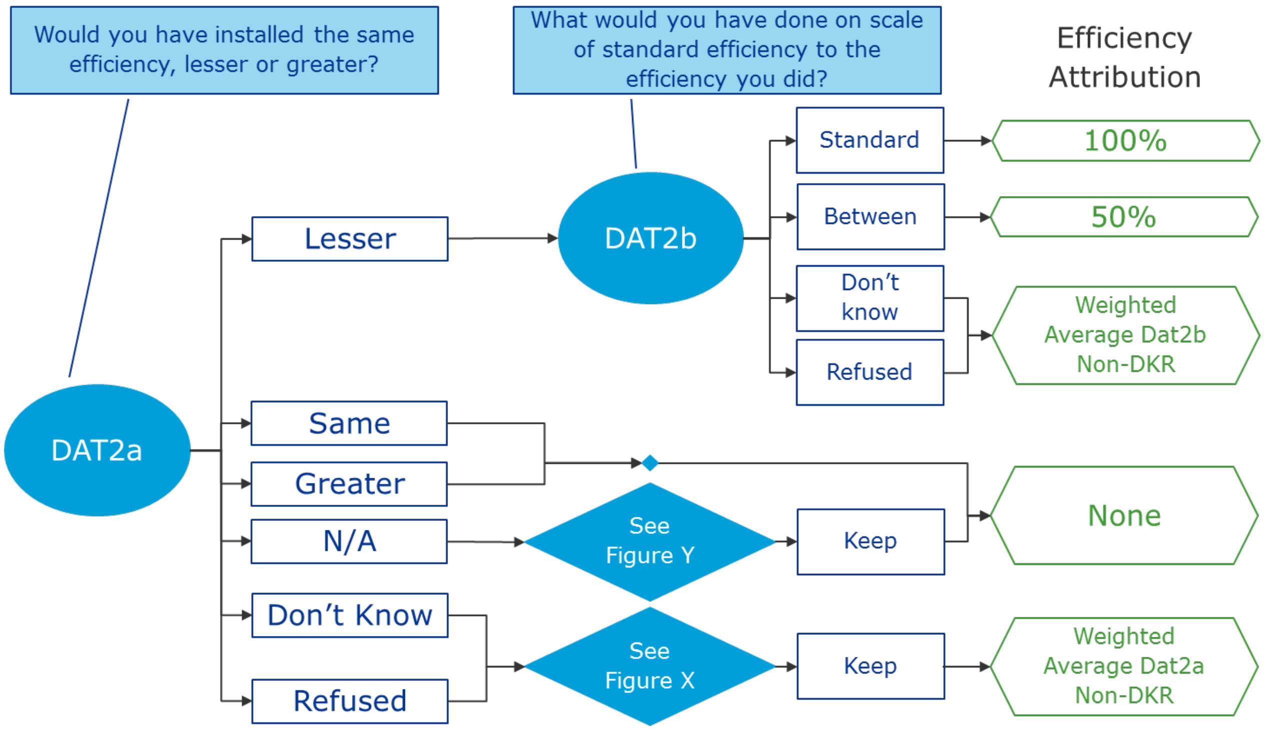
Efficiency

Efficiency Attribution, A_E, gives the program credit for increasing the efficiency of a measure above what would have been installed in the absence of the program. The two relevant questions are DAT2a and DAT2b.

- DAT2a: “Without <the program>, would you have installed the same efficiency as what you installed, lower efficiency, or higher efficiency?”
- DAT2b: “Without <the program>, would you have installed <measure> that was “standard efficiency on the market at that time,” or “between standard efficiency and the efficiency that you installed?” (DAT2b is only asked if DAT2a is “Lesser.”)

The program receives nonzero Efficiency Attribution if the respondent indicates that they would have installed a less efficient measure without the influence of the program. The magnitude of the Efficiency Attribution is determined based on the answer to DAT2b, as shown in **Error! Reference source not found.**. **Error! Reference source not found.** shows the corresponding decision tree for DAT2a and DAT2b.

Figure 26: Decision Tree for Efficiency Attribution



If the respondent answers DAT2a with Greater or Same then the survey skips to the next section and there is zero Efficiency Attribution. If efficiency is not applicable to this measure but quantity is applicable and the measure would have been installed anyway then the survey skips to the next section and the Efficiency Attribution will not affect the Simple Program Attribution. If the respondent answers DAT2a with Don't Know or Refused but does provide answers to inform the Quantity Attribution and Acceleration Period then the measure is assigned the average Efficiency Attribution for all measures in the same measure group.

Table 49: Efficiency Attribution Assignments

Coarse Cut (DAT2a) (what efficiency would have been implemented absent the program)	Finer Cut (DAT2b)	Efficiency Attribution
Same	NA	0%
Lower	Standard efficiency or according to code	100%
	Between standard efficiency and the efficiency that was installed	50%
	Don't know/refused	Weighted average of above cases for primary domain
Greater	NA	0%
Don't know/refused	NA	Weighted average of all respondents for primary domain

Quantity

Quantity Attribution, A_Q, gives the program credit for increasing the quantity of a measure above what would have been installed in the absence of the program. The two relevant questions are DAT3a and DAT3b.

- DAT3a: “Without <the program>, how different would the <number/size> of the <equipment type> have been? Would you say you would have installed the same amount, less, more, or not have installed anything?”
- DAT3b: “By what percentage did you change the amount of <equipment type> installed because of <the program>?” (DAT3b is only asked if DAT3a is “Less” or “More.”)

The program receives nonzero Quantity Attribution if the respondent indicates that they would have installed less of the measure or a smaller measure without the influence of the program (for example: “I would have replaced as many doors”_. The program also receives nonzero Quantity Attribution if the respondent indicates that they would have installed more of the measure or a larger measure without the influence of the program (for example: “I would have installed a bigger furnace, but I through the program I learned it was unnecessary”). The latter case covers situations where the program effect was in “right sizing” the measure. The magnitude of the Quantity Attribution is determined based on the answer to DAT3b, as shown in **Error! Reference source not found.**. **Error! Reference source not found.** shows a decision tree for DAT3a and DAT3b.

Figure 27: Decision Tree for Quantity Attribution

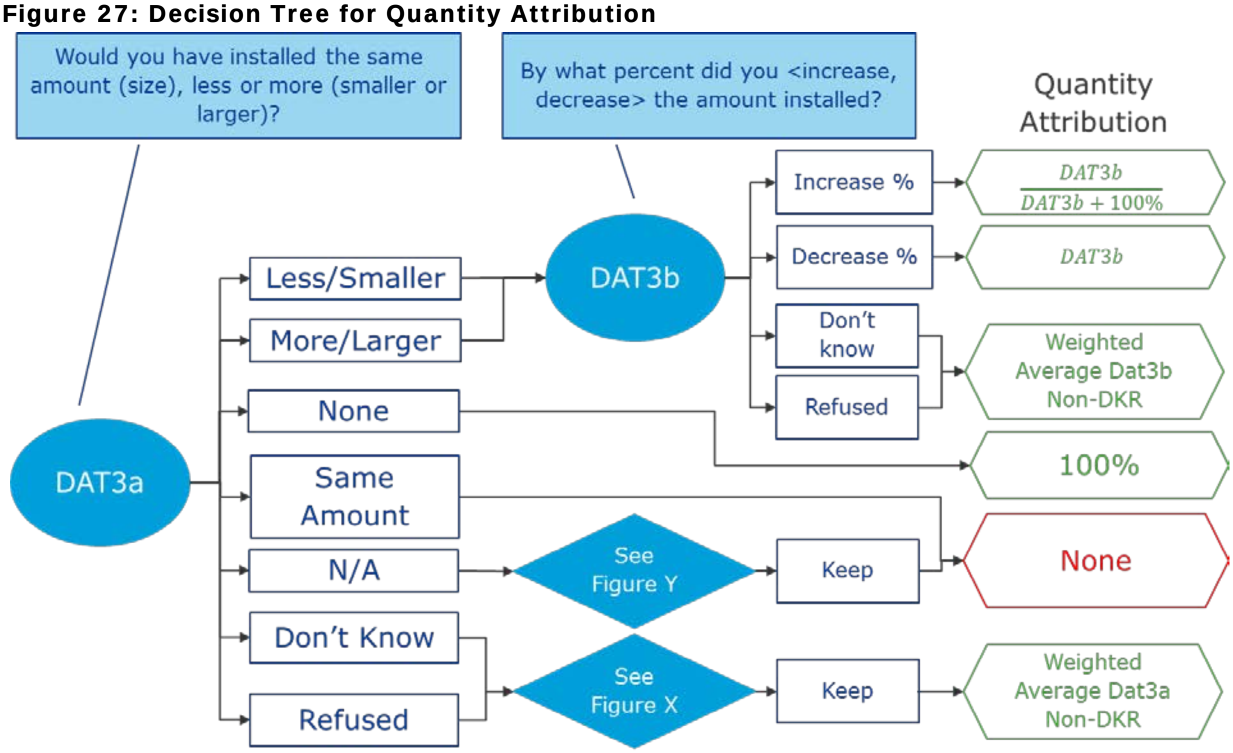


Table 50: Quantity Attribution Assignments

Coarse Cut (DAT3a) (How much equipment would have been replaced absent the program)	Finer Cut (DAT3b)	Quantity Attribution
Same	N/A	0%
Less/Smaller	ΔQ	AQ = ΔQ / (ΔQ + 100%)
	Don't know/refused	Weighted average of "less" cases for primary domain
More/Larger (right sizing)	ΔQ	AQ = ΔQ
	Don't know/refused	Weighted average of "more" cases for primary domain
None	N/A	100%
Don't know/refused	N/A	Weighted average of all respondents for primary domain

If the respondent would have installed a smaller measure without the program then the Quantity Attribution is calculated as:

A_Q = Inc / (Inc + 100%)

where

Inc = percentage change in quantity because of the program.

If the respondent would have installed a larger measure without the program, then the Quantity Attribution is calculated as:

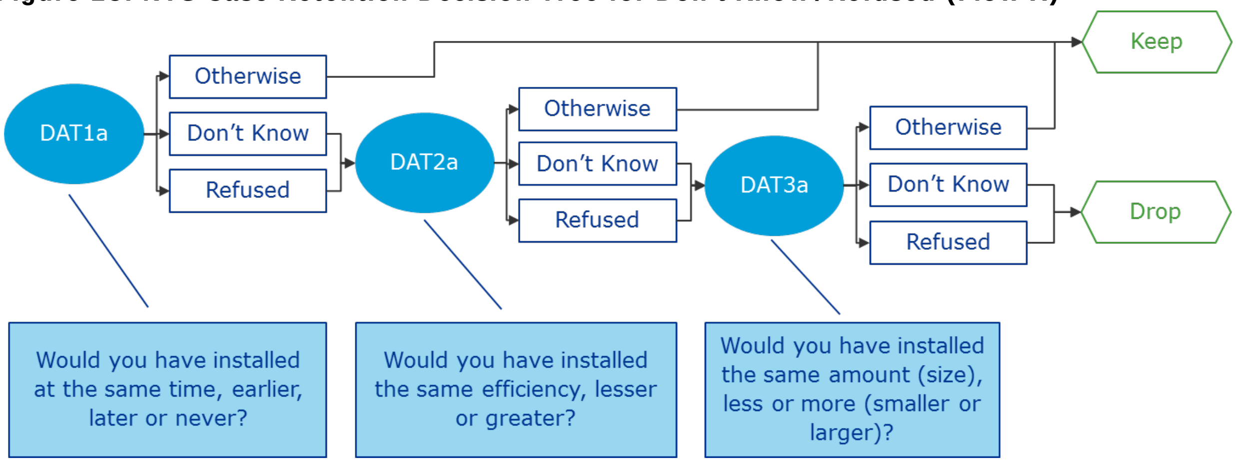
A_Q = Inc.

If the respondent answers DAT3a with Same Amount or None then the survey skips to the next section and there is zero Quantity Attribution. If quantity is not applicable to this measure but efficiency is applicable and the measure would have been installed anyway then the survey skips to the next section and the Quantity Attribution will not affect the Simple Program Attribution. If the respondent answers DAT3a or DAT3b with Don't Know or Refused but does provide answers to inform the Efficiency Attribution and Acceleration Period then the measure is assigned the average Quantity Effect for all measures in the same measure group.

What if they "Don't Know" or "Refuse?"

Some respondents are unable or unwilling to answer the relevant questions in the survey attribution sequence. If a participant is unable or unwilling to answer *all* of the attribution questions then the participant is dropped from the attribution analysis. However, the respondent information will still be included as part of the installation rate and the VGI. **Error! Reference source not found.** shows a decision tree that indicates the relationship between the question responses and how they affect the attribution. If a measure goes to the "Keep" decision then the ultimate resolution of each effect is shown in **Error! Reference source not found.**.

Figure 28: NTG Case Retention Decision Tree for Don't Know / Refused (Flow X)

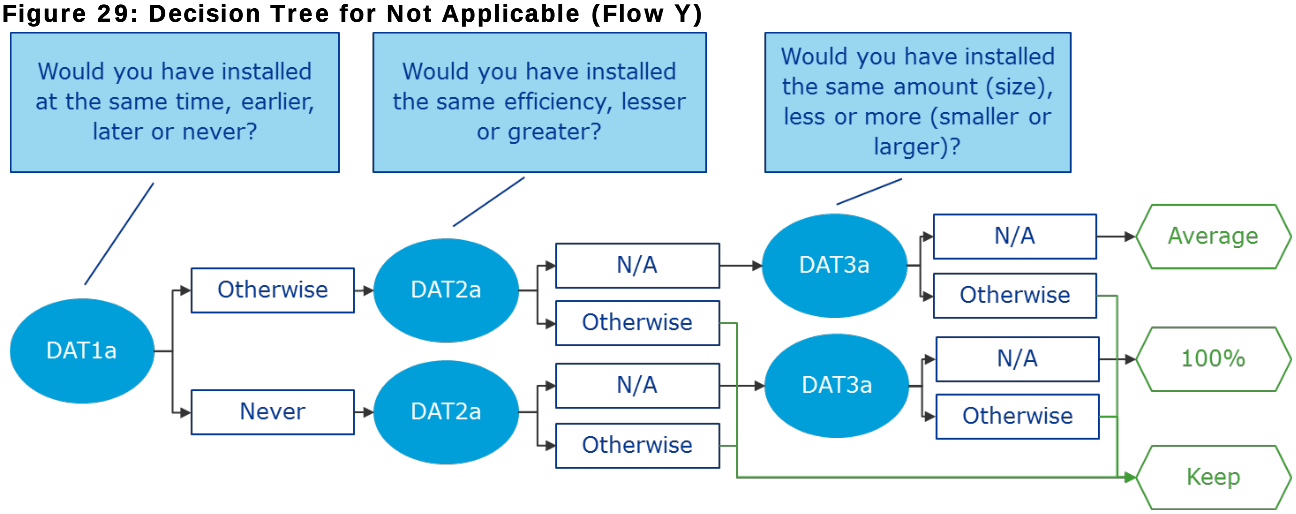


When Efficiency and Quantity Don't Apply

Quantity and efficiency questions do not apply to all measures. Efficiency questions do not apply if the equipment type is inherently an efficiency improvement; that is, the "standard efficiency" baseline would be not to install anything. Variable frequency drives (VFDs) or heat recovery systems are examples. Quantity questions do not apply when varying quantity or size does not make sense in the context of the measure.

Error! Reference source not found. shows a decision tree that indicates the relationship between the question responses and how they affect attribution. If a respondent indicates that a measure would never have been installed without the program and the DAT2a and DAT3a questions do not apply then the attribution is 100%. If the respondent would have installed the project at the same time, earlier, or later and the DAT2a and DAT3a questions do not apply then the measure is assigned the average savings-weighted attribution across all measures in that measure group.

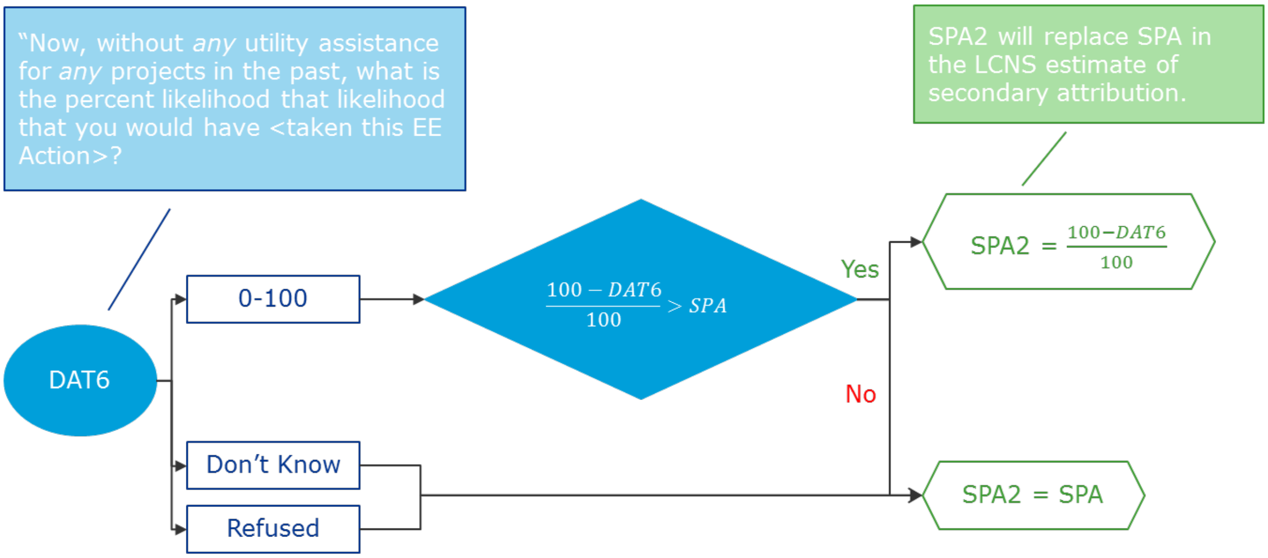
Figure 29: Decision Tree for Not Applicable (Flow Y)



Secondary Attribution

Secondary attribution, the longer-term effect of the program on participant decision making will be assessed based on a single question. That asks the respondent about the likelihood of the project given all program assistance for all projects since the programs were started. The greater of this score and the primary SPA will be used as SPA in calculating the LCNS based on all program efforts, not just those focused on this project.

Figure 30: Secondary Attribution Scoring

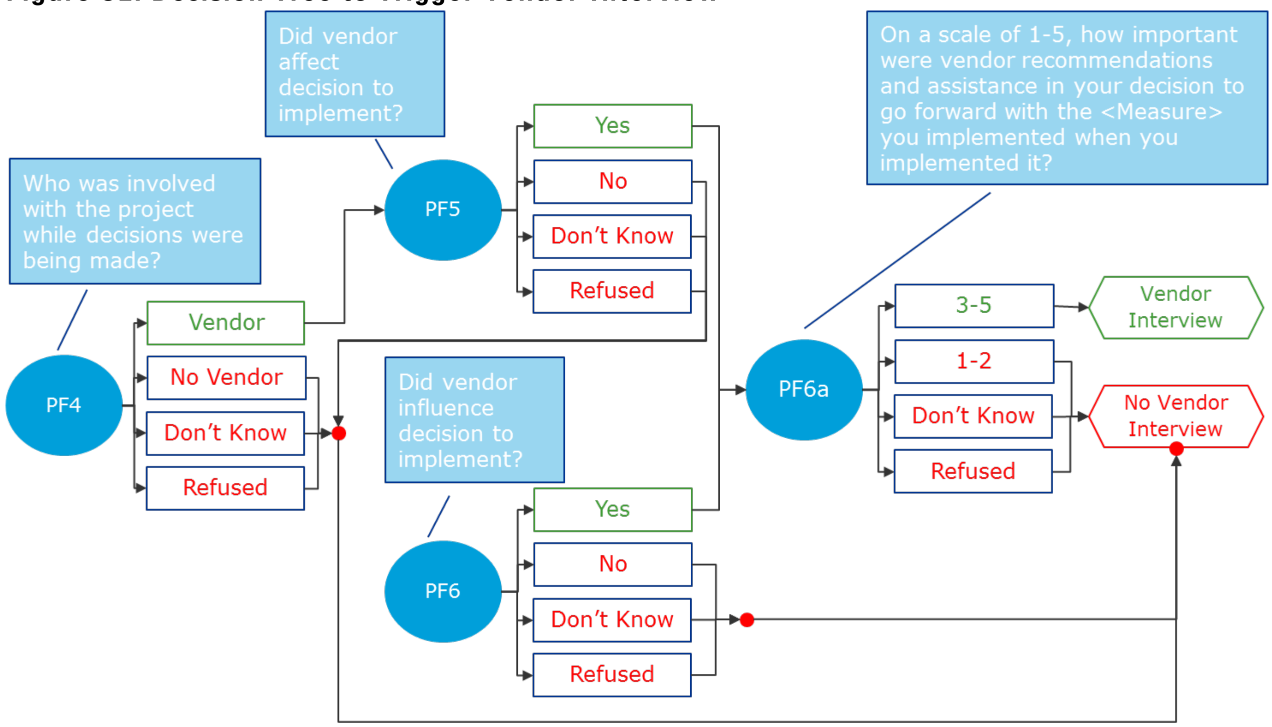


Incorporating Vendor Effect

DNV GL will take two steps to determine when a vendor survey is necessary to supplement the participant survey. They are:

1. When we request project documentation and site contact information for each sampled project we will also ask the utilities to provide vendor contact information for projects with vendor involvement.
2. Each survey completed with a participant is reviewed to determine the effect the supplier had on the participant's decision to install a given measure relative to the program's effect. If a participant indicates that the program did not influence their decision to install high efficiency equipment but the vendor did have substantial influence, then we will complete a survey with the vendor. The decision tree is shown in **Error! Reference source not found.**.

Figure 31. Decision Tree to Trigger Vendor Interview



For measures with both participant surveys and vendor surveys, the analysis will produce two separate attribution values. The first reflects the influence that the program had on the participant's decision to install the measure. The second reflects the influence that the program had on the vendor's business practices and therefore their ability to sell the measure. We choose the higher of the two values as the final program attribution for that measure. That is, if either the vendor or the customer indicates that the program influenced the decision to install the measure, we conclude that the program influenced the decision. In the event that a vendor interview is triggered, but is either not completed or results in an inconclusive vendor score, vendor attribution for the measure will be the average attribution of all completed vendors within the evaluation program.

The vendor attribution scoring method will be included with the vender interview guide.

Quality control by interviewers and analysts

Each of the components of attribution, Timing (DAT1a/ DAT1a_O/DAT1b), Efficiency (DAT2a/DAT2a_O/DAT2b) and Quantity (DAT3a/ DAT3a_O/DAT3b), have a question sequence that follows the same pattern:

DATXa.	What would you have done without the program?
DATXa_O.	Why do you say that?
DATXb.	<If DATXa=program effect> How different would the project have been?

Quality control for each component of attribution consists of comparing the final component attribution score (t, e, q) to the open ended response for the "DATXa_O. Why do you say that?" question.

1. Interviewers are trained to probe if the response to the open ended question is inconsistent with the scored response to DATXa.
2. During the analysis phase, the analyst will put measures into three bins: full attribution, partial attribution and full free rider for each component. The analyst works bin at a time to compare each verbatim open ended response to the score for the attribution component. Assessing verbatim responses by bin reduces analyst error and speeds the review. If an open ended response appears inconsistent with the score received, the case is elevated to PM review.

Overall attribution scores are compared to the DAT0 score and assesses for consistency. A high attribution score from the TEQ questions should correspond to a "somewhat unlikely" or" very unlikely" to implement response to DAT0. Inconsistent scores are referred to PM review.

Overall attribution scores are also assessed for consistency with the DAT4 verbatim, by bins as described for the QC of the component scores. Inconsistent scores are referred to PM (Ben Jones) review.

Non-Zero attribution scores are also assessed for consistency with the responses to PF8 and PF9. Any non-zero score that also has a response of "after making decision" or "after installing" is considered inconsistent and referred to PM (Ben Jones) review.

The overall attribution score will also be compared to DAT6 (the secondary attribution question). In theory, DAT6 should be equal to or greater to the overall attribution score for all measures, but because the question is a scalar 1-10 and the primary attribution is scored by asking about influence on specific aspects of the project inconsistencies are expected. For QC, all instances where the secondary attribution is more than 20% lower than overall primary attribution will trigger a PM review of.

Quality control PM Review

Analysts are instructed to have a low bar ("when in doubt flag for review"), most of the measures flagged for PM review result in no change. For each site that has a measure flagged for PM review, the PM (Ben Jones) will review the full survey, including all measures and responses. The PM may also follow up with the interviewer to better understand the combination of responses. If the PM determines that the flagged score (whether of a component or overall) is not clearly contradicted by the overall story told by the respondent throughout the interview, the PM makes no change. If the flagged score is clearly contradicted (approximately one percent of cases in DNV GL's experience), the PM decides among three options:

1. drop the measure from the sample (for very muddled responses, much more common with CATI than IDI)
2. replace the inconsistent response with a "Don't Know" (effectively using the average if it is clear that there should be some attribution for the component, but unclear how much)
3. adjust the flagged score to more accurately reflect the intent of the respondent (employed in cases where there is overwhelming evidence of intent, for instance the open-ended response says clearly what the score should be)

For all adjusted scores, project sponsor (Tammy Kuiken) approval is required.

An appendix in the FR report will provide information on the QC process including:

1. (anonymized) verbatim responses by attribution bin for each component (DATXa_O) and overall (DAT4)
2. how many scores were adjusted and how
3. cross-tabs of
    a. DAT0 response versus overall attribution bin
    b. DAT6 response versus overall attribution bin
    c. PF8 responses by overall attribution bin
    d. PF9 responses by overall attribution bin

APPENDIX D SPILLOVER METHODOLOGY

The spillover analysis will provide separate estimates of spillover for inside-like, inside-unlike, outside-like, and outside-unlike spillover. Each of the estimates will be generated based on ratio estimation relative to the program measure savings.

Spillover "refers to effects of customers that adopt energy efficiency measures because they are influenced by a utility's program-related information and marketing efforts, but do not actually participate in the program."[42] As in many jurisdictions, Ontario's Demand-Side Management Guidelines recognize the importance of spillover in determining program benefits, and also require "comprehensive and convincing empirical evidence" to support any program spillover claim.

Key challenges to providing convincing quantified evidence of spillover for a particular customer include:

- Determining that a particular subsequent action was due to the influence of the program
- Confirming that the action was not taken as part of the original or another program, hence already counted by the program
- Quantifying the savings associated with confirmed spillover actions.

DNV GL's proposed approach provides a high level of rigor to address each of these issues.

- We determine program influence using participant surveys that start with the framing used for our free ridership questions. This framing helps ensure more meaningful responses to questions of the influence of the experience with the program in implementing the original measure on subsequent actions. As for the free rider surveys, obtaining the right respondents is also essential to obtaining meaningful responses to these questions.

- We confirm that the actions tentatively identified as spillover were not already counted by another program by cross-checking tracking data bases. Also critical to separation of spillover from program-claimed savings is understanding what savings if any are claimed by the programs for facilitation support such as opportunity identification, feasibility studies, audits, and related continuous improvement program engagement.

- We quantify the savings for confirmed spillover actions by collecting engineering specifications and calculating associated savings. This approach gives more accurate results than asking customers to estimate the magnitude of spillover savings relative to the original measure.

Thus, our participant spillover methodology addresses the following key issues:

- *Locating the right decision-maker* - Large commercial and industrial companies have multiple decision-makers and it is often difficult to find someone who is familiar with both the tracked program-influenced measure and the spillover measure. Employee turnover can also complicate this. Our approaches to ensure appropriate respondents are discussed above.

- *Avoiding double-counting* – Companies that received financial incentives from an energy efficiency program for one measure are likely to seek these incentives for future measures. Hence it is important to get the program's latest tracking data to make sure that a potential participant spillover measure did not receive program support.

[42] Ontario Energy Board *Demand Side Management Guidelines For Natural Gas Utilities,* EB-2008-0346, June 2011, Chapter 7.

- *Estimating program attribution for potential spillover measures* - A common way of assessing participant spillover is to ask how much the participant's experience with the tracked program-influenced measure influenced their decision to implement measures that are candidates for spillover attribution. It is difficult to turn this "fuzzy" assessment of program influence into a more concrete attribution factor necessary for attributing a certain quantity of m^3 from the spillover measure to the program.

- *Estimating the energy savings for the participant spillover measures.* Because spillover measures occurred outside the program, evaluators do not have access to the same information about the size, type, and quantity of the implemented energy-efficient measures that they would find in a program tracking database.

Our approach to these issues is described in more detail below. This approach is based on one we used successfully in Wisconsin C&I programs over many years.

Understanding Energy-Related Standard Practices

The first objective of the survey will be to find out whether the participant's company or organization had installed any energy-efficient equipment or made any energy-efficient changes in operation or maintenance (O&M) procedures after the implementation of the tracked project. But before doing that we will collect some information about the company or organization's energy-related decision-making process. We will ask the participants a series of questions about:

- Who in their company makes decisions about equipment replacement and retrofit projects;
- What information sources are used in making these decisions; and
- Possible barriers to energy efficiency implementation.

This information will be valuable for a number of different reasons. First, it should help program implementers devise strategies for increasing program awareness and mitigating barriers to project implementation, especially for participants who did not identify any subsequent energy-efficient projects after the tracked project. Second, by shedding light on the project decision-making process, it should help the evaluators make better judgments about assigning program attribution to a given project. Finally, it should make the survey appear less peremptory for those who did not report any new energy-efficient projects after the tracked projects, since otherwise their survey would be terminated fairly quickly.

After we collect this information about participant energy practices, we will ask the participants whether their company/organization had installed any energy-efficient projects after the installation of the tracked project. If the participants report no subsequent actions, we will terminate the survey since there is no participant spillover to be measured. If they do identify subsequent projects, then we will collect some basic information about the project including:

- The approximate year of the project;
- The geographic location of the project (e.g. city);
- The types of energy-efficient measures installed or energy-efficient O&M practices implemented; and
- Whether the tracked project and the subsequent project were in the same facility or not (needed for the calculation of inside vs. outside spillover).

Because this information will be collected by CATI program surveyors who do not have an energy background, we will not ask them to try to collect too detailed information about the energy-efficient project.

It just needs to be detailed enough to allow the evaluators to make a reasonable match with any projects in the program tracking data.

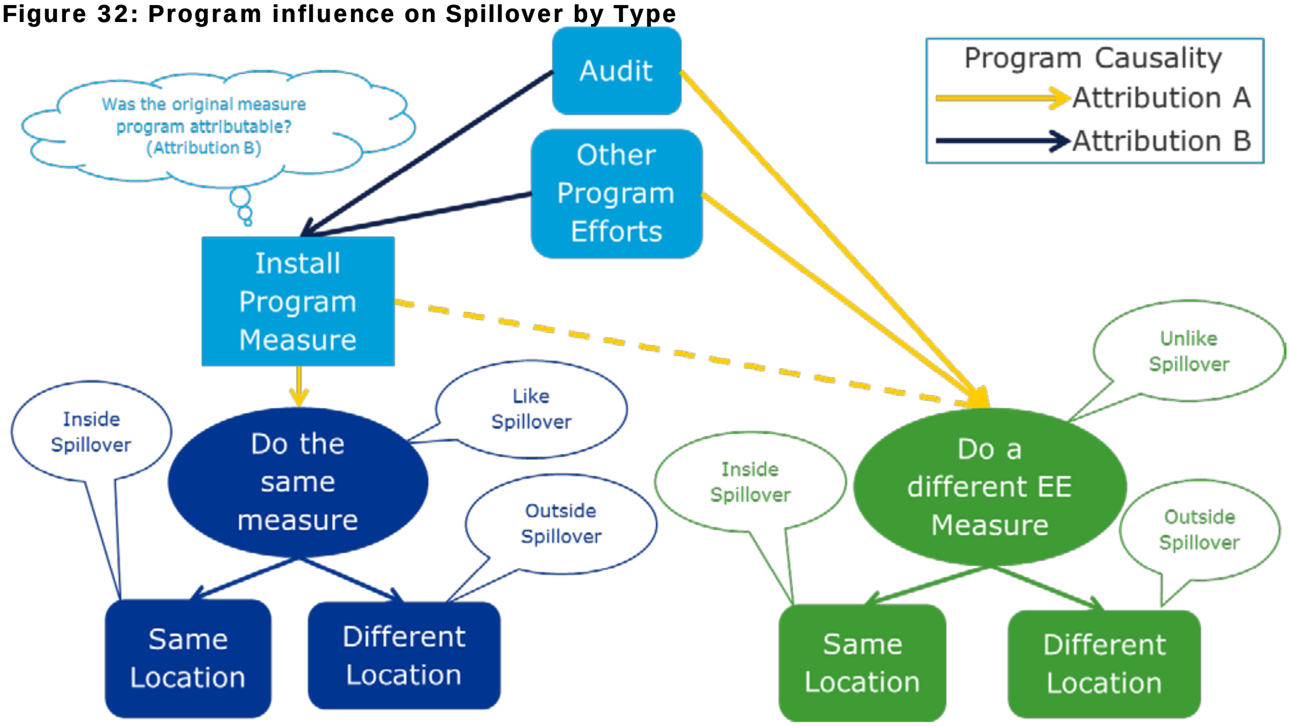
Calculating Program Attribution for Candidate Spillover Actions

The next stage of the survey will focus on program attribution. Our method only awards spillover energy savings if two criteria are met:

1. The potential spillover project is at least partially attributable to the participant's experience with the program in implementing the earlier tracked project (Attribution Factor A).
2. For like spillover, the original tracked project is at least partially attributable to the program (Attribution Factor B). For unlike spillover, Attribution B will apply if the respondent indicates that the original program measure (separate from other program efforts) was a factor in their decision.

Figure 32 shows how program causality ties to different types of spillover. Attribution B applies to like spillover in all cases, while for unlike spillover attribution B applies to the spillover only if the original program measure was part of the program influence that led to the spillover measure being implemented.

Figure 32: Program influence on Spillover by Type



If a measure met these two criteria, we assign it spillover savings according to the following formula.

(Spillover Savings) = (the measure's savings) X (Attribution Factor A) X (Attribution Factor B).

We apply both Attribution Factor A and Attribution Factor B because if the program had no influence on the original tracked project, the program should not get credit for any additional measure installations resulting from that tracked project. To reduce respondent fatigue, Attribution Factor A will be asked in the CATI survey, while Attribution B will only be asked in the Engineering follow up IDI. If Attribution A is zero we will not

To determine attribution factor B we will use the FR question battery already described in this SOW. For Attribution factor A we will use a scoring method that will be triggered off the question, "If you had not made the earlier energy-efficiency improvements I just listed, how likely would you have been to make this additional energy efficiency improvement?" The scoring method, which we used in Wisconsin for many years, is shown in Table 51. If the participant said they were very likely to have made the additional energy efficiency improvement without the program, then we will terminate the survey since there will be no participant spillover to be measured. If the subsequent measure is fully or partially attributable, then for unlike spillover a follow up question will be administered to assess whether Attribution B is applicable.

Table 51: Program Attribution for Subsequent Measures

If had not made tracked program-influenced energy efficiency improvement, reported likelihood of making subsequent energy efficiency improvement		Assigned Attribution Factor A
1	Not likely at all	1.00
2	Not very likely	0.90
3	Somewhat likely	0.55
4	Very likely	0.00

The reason we use a different method for Attribution Factor A than for Attribution Factor B is that the character of influence is different. For the program's influence on the tracked project (Attribution Factor B) financial incentives usually account for much of the influence in terms of reducing payback periods and therefore we want to measure things like acceleration effects. However, with participant spillover the influence is less tangible and more likely to be general positive experience with a new energy-efficient technology and the energy savings it produces. We believe that using a Likert scale question will better capture the less tangible character of this type of influence.

Avoiding Double Counting of Energy Savings

Once a participant has identified a subsequent project that is attributable – e.g. one where Attribution Factor A and Attribution Factor B are both greater than zero -- then we will conduct some additional checks to insure that the subsequent project is not also a tracked project. Some of these checks will occur in the survey itself. For example, we will ask the participants if they recall receiving financial incentives from an energy efficiency program for the subsequent projects. We will also examine the program tracking data to make sure that the subsequent project is not in the tracking program data for future years. For example, if we interview a 2013 participant and they identify a subsequent project in 2014 we will look at the 2014-2015 program tracking data (we will look at both program years in case their memory of the project timing was faulty) to see if we can find that project. If we do find the subsequent project in program tracking data, then we will remove that project as a candidate for spillover energy savings since the savings for that project has already been claimed by the program.

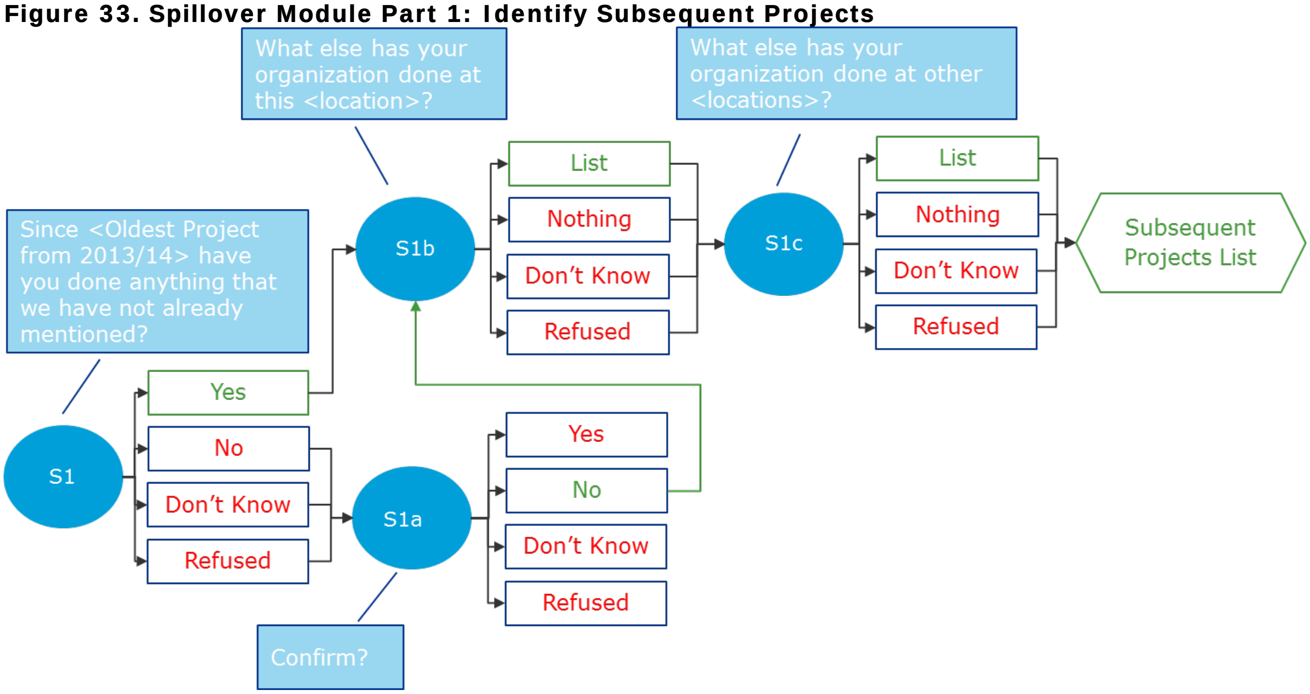
Estimating Energy Savings for Participant Spillover Measures

Once a project has been identified as having spillover energy savings (it is program attributable and we could not locate it in the program tracking data) the final step will be to estimate its energy savings. To estimate the annual energy savings for participant spillover measures, we plan to have engineers conduct follow-up interviews with the persons identified in the CATI surveys as being most familiar with the spillover projects. The engineers will have some basic project information collected from the CATI survey as well as some information about deemed savings algorithms for that measure which will allow them to prepare ahead of time the types of questions they will need to ask (e.g., about baseline measures, hours-of-use, etc.). Once they have conducted the interview and collected the necessary information they will calculate the first year savings and EUL for the measure. If a deemed savings algorithm exists for that measure they will use that as a default. If none exists then they will use their best professional judgment to estimate the energy savings. This process will work equally well for both like and unlike spillover.

Spillover Decision Trees

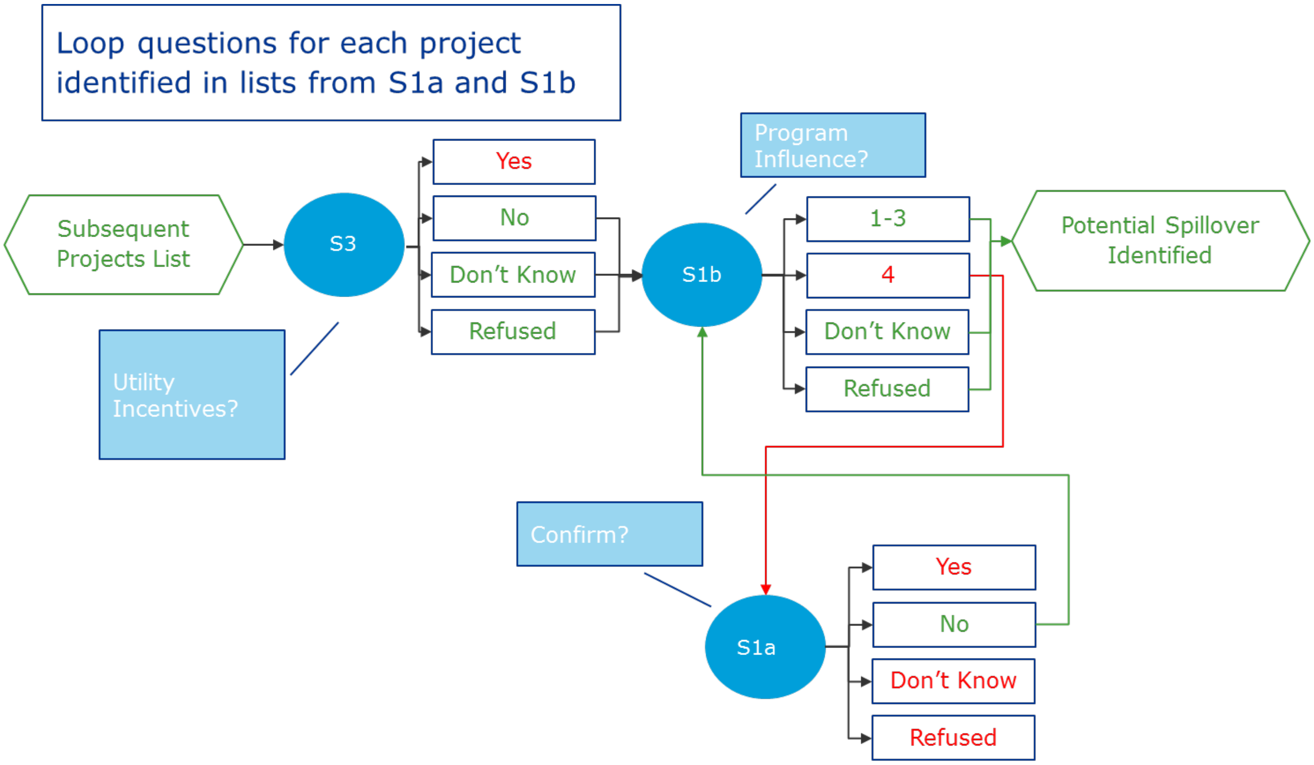
The initial participant IDI and participant CATI each include a spillover module that produces a list of potential spillover projects for each participant. The first part of the module (Figure 33) generates a list of changes to energy using equipment at the same location as the original measure and another list of changes to equipment at other locations.

Figure 33. Spillover Module Part 1: Identify Subsequent Projects



The second part of the module (Figure 34) loops through the list of subsequent projects to eliminate projects that received utility incentives and to establish program influence. The projects identified that were program influenced are referred to as potential spillover and will receive a follow up engineering interview to quantify savings.

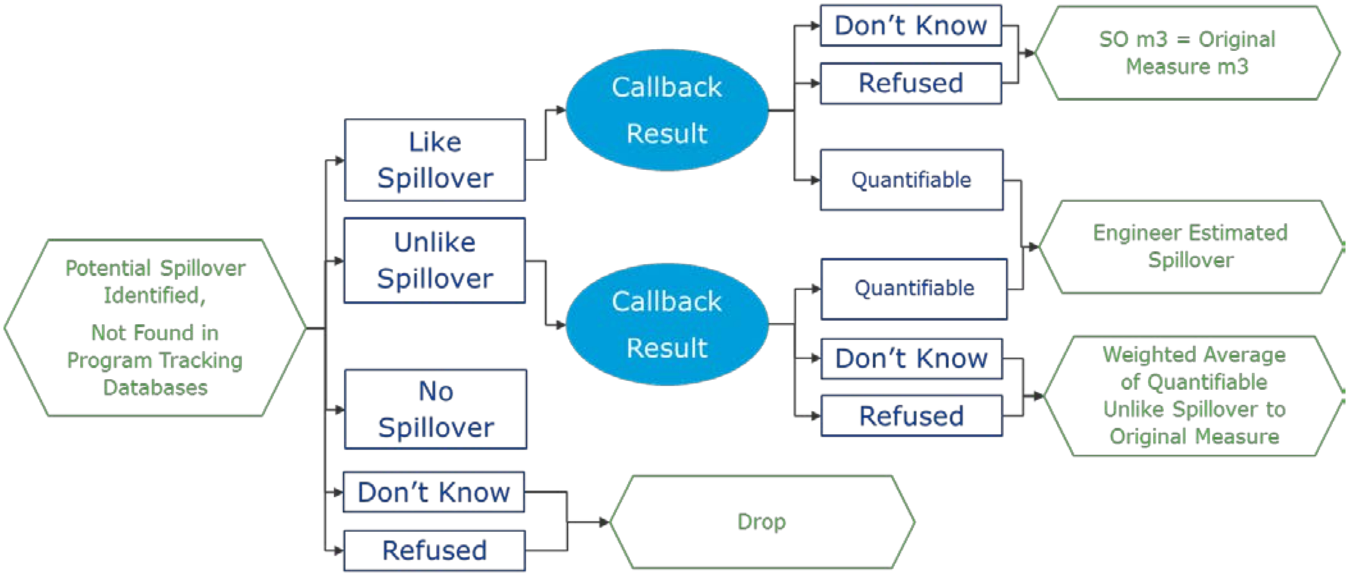
Figure 34. Spillover Module Part 2: Subsequent Project loop



72

Potential spillover projects that are not found in program tracking databases will receive a call from a DNV GL engineer. If the customer refuses the interview or the evaluation engineer is not able to find a contact who can answer technical questions, the spillover will be quantified in one of two ways (Figure 35). If the project is like spillover we will use the savings of the original program measure as the basis for the savings estimate. If the project is unlike spillover we will use the average of other sites with unlike spillover for the estimate.

Figure 35. Spillover Callback High Level Process



APPENDIX E SAMPLE EXPANSION AND RATIO ESTIMATION

Sample Weights

This appendix describes how we calculate the sample weights for each stratum. In lay terms the weight is simply the number of units in the sample frame (N) divided by the number of completed units in the sample (n). The interpretation of the weight is that each completed sample unit represents N/n units in the population (sample frame).

Notation:

N_x = number of units of analysis in stratum X

n_x = number of completed sample units of analysis in stratum X

The weight W_x is calculated as

$$W_x = N_x / n_x$$

We can understand the weight as meaning the response for one sampled unit in stratum X is representative of W_x units in the population. Table 2 shows a simple example. In the example we completed 2 surveys with participants in the "North" and 10 surveys with participants in the "South." The weight for the "Northerners" is greater than that of the "Southerners," but because we completed more surveys with "Southerners" the combined weight of the "South" will be in proportion to its share of the population (both the population and sum of weights is 20).

Table 2: Example Sample Weights

Stratum Definition	Sample Frame (N)	Sample Completes (n)	Weight (W)	Interpretation
North	10	2	5 = 10/2	Each response represents 5 Northern participants
South	20	10	2 = 20/10	Each response represents 2 Southern participants

Without sample weights, the data collected from the "North" would be 17 percent (2/12) of the final result, while with weights, the "North" is 33 percent (10/30). The un-weighted result would be less accurate than the weighted result if the measured value differs along North/South lines. For example if the "North" is more conservative than the "South" then political surveys without sample weights would end up with inaccurate results. If responding to surveys is negatively correlated with conservatism, then the weights help correct for the systemic bias in response rates.

The sample weight associated with an observation is consistent regardless of the segmentation of the data that we report by (reporting domains). This means that we can segment the data multiple ways in the report, with the final overall results consistent no matter the domain.

Special Cases

There are some special cases where the sample weight for a project needs to be set to 1 in order to use the data collected without biasing the result. Our sample design targets measures within a site and sample weights are developed at that level as well. When we collect data from a customer we will collect data on all of a customer's measures in a single IDI or site visit. This maximizes the data collected on each customer contact, but requires special handling to ensure that extra data collected does not bias the sample. To eliminate the potential bias of over representing multiple measure sites, we first identify units that were completed as an add-on when another measure was selected for a site.

For each stratum in our sample design the units are randomly ordered for selection in a list. If seven units are targeted for the stratum then the first seven units on the list are the primary sample and the rest of the list comprises the full backup sample (when we request project documentation we will restrict the backup sample for the request in order to reduce burden on utility staff). If a site has two measures in different strata and one is selected in the primary sample, we will request documents on both measures and ask about both, regardless of whether the second measure is in the primary or backup sample in its stratum. After collecting data on both measures we will assess whether the second measure was selected in its stratum based on how far down the list we had to go to complete our target. If the second measure's spot on the list was selected, then the measure will be counted as a normal complete and included in the stratum's N/n weight calculation. If the measure's spot on the list did not come up, the data collected for the measure will be used, but the measure will not be included in the N/n weight for its strata. Instead it will be given a weight of 1 so that it represents itself and no other measures. For variance estimates, the measure will remain in its sampled stratum.

Table 52 provides an example. Both site A and Site B were had measures in Stratum X selected in the sample. Each responded to our interview. Both sites also had a measure in Stratum Y. The evaluation completed data collection for both measures for each site. Due to where each of the sites' second measures were on the original priority list in stratum Y, the second measure for each site received different weights despite being in the same stratum.

Table 52: Determining non-randomly selected measures

Strata	Priority	Site	Measure	Survey Disposition	Selection Type	Weight
X	1	**A**	**A1**	**Complete**	**Random**	**3 / 2**
X	2	**B**	**B1**	**Complete**	**Random**	**3 / 2**
X	3	C	C1	live		
Y	1	D	D1	Complete	Random	8/3
Y	2	E	E1	Refused		
Y	3	**A**	**A2**	**Complete**	**Random**	**8 / 3**
Y	4	F	F1	Complete	Random	8/3
Y	5	G	G1	live		
Y	6	**B**	**B2**	**Complete**	**Not Random**	**1 / 1**
Y	7	H	H1	live		
Y	8	I	I1	live		
Y	9	J	J1	live		

The measures in Stratum X each were selected randomly. Measure A1 was first on the priority list and measure B1 was second. Because both A1 and B1 were completed and the target was 2 for the strata, site C was not called. Because site C was not called, measure C1 had a final survey disposition of "live." In the case of stratum X, there were 3 measures and 2 were completed. This resulted in a sample weight of 3/2 for each of the two completed measures.

In stratum Y four measures were completed. In this example the target for the stratum was achieved prior to calling site G. The evaluation attempted data collection for the first 4 measures on the list. Site E refused the survey or otherwise did not respond. Sites D, A, F and G completed the survey, but B did not come up in the priority list until after site G (the first "live" site in the list). In this case measure B2 was not selected randomly and needs to be treated as a special case. Measure B2 is removed from the stratum Y weight calculation, so the three measures that were completed receive a weight of 8/3 (once measure B3 is removed there are eight measures in the frame, and 3 completed measures). Measure B2 receives a weight of 1.

Ratio Estimation

The calculation of the adjustment factors for tracking system gross and net savings uses appropriate case weights corresponding to the sampling rate as discussed above. The energy saving estimates (tracking savings, installed savings, verified savings or net savings) of the sampled units (measures, projects, sites) are present in both the numerator and the denominator of the ratios, when combined with the sample weights the ratio estimation method produces unbiased, savings weighted adjustment factors.

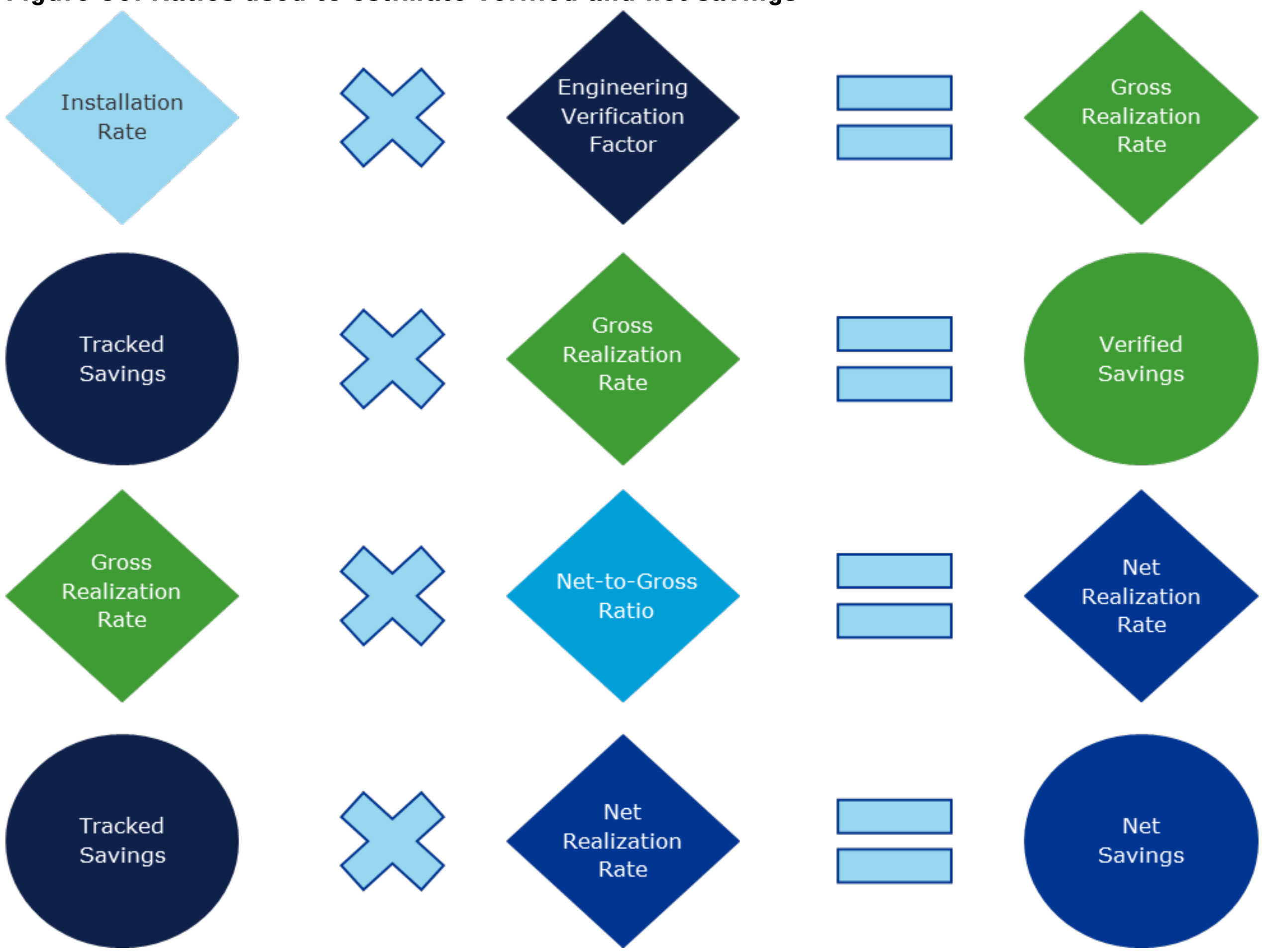
Collecting data on verified and net savings for the same set of measures provides a more accurate estimate of net savings. Integrating the two allows the evaluation to calculate net savings for a measure as a function of verified savings rather than tracking savings. This means that projects carry the weight of their specific verified savings in the net-to-gross ratio rather than tracking savings or a broader estimate of verified savings. Large verification adjustments can have a large effect on the relative weight of specific projects in the NTG.

For an individual measure:

- Installed savings are a function of the tracking savings. When the measure is installed the installed savings equal tracking savings and when the measure is not installed, then installed savings are zero.
- Verified savings are calculated independent of the tracking savings by evaluation engineers using the best available methods and information.
- Net savings are a function of verified savings. Attribution for the measure multiplied times verified savings plus spillover savings associated with the measure.

Individual measure results are expanded to the estimate population savings (circles) using ratios (diamonds), as shown in **Error! Reference source not found.**. Ratios are applied for each of the primary reporting domains and then summed to calculate the total for the program overall.

Figure 36: Ratios used to estimate verified and net savings



Two general ratio calculation approaches are employed: directly calculated and combined. The description of the process is easiest to understand through an example. The example below has three directly calculated adjustment factors: the installation rate, the engineering adjustment, and the net-to-gross factor. Each of these is calculated as a ratio estimator over the sample of interest (Cochran, 1977, p.165). The formulas for these factors are given below.

Notation: The following terms are used in calculating the adjustment factors:

G_{Tj} = tracking estimate of gross savings for measure *j*

G_{Ij} = tracking estimate of gross savings for measure *j*, adjusted for non-installation

G_{Vj} = engineer verified estimate of gross savings for measure *j*,

N_{Vj} = Net verified estimate of gross savings for measure *j*,

W_{Vj} = weighting factor for measure j used to expand the CPSV sample to the full population

W_{Nj} = weighting factor for measure j used to expand the FR sample to the full population

The installation rate R_I is calculated using the CPSV sample as

$$R_I = \frac{\sum_{j \varepsilon A} G_{Ij} w_{Vj}}{\sum_{j \varepsilon A} G_{Tj} w_{Vj}}$$

The Engineering Adjustment R_E is calculated from the CPSV sample as

$$R_E = \frac{\sum_{j \varepsilon V} G_{Ej} w_{Vj}}{\sum_{j \varepsilon V} G_{Ij} w_{Vj}}$$

The Attribution ratio R_A is calculated from the FR sample as[43]

$$R_A = \frac{\sum_{j \varepsilon N} N_{Vj} w_{Nj}}{\sum_{j \varepsilon N} G_{Vj} w_{Nj}}$$

The procedure used for calculating ratio estimation by domains provides the correct standard error of the estimate for each domain and overall. The procedure also takes into account defined clusters of observations (customers) and stratification. The standard error is calculated using two methods.

The first method recognizes the sample as drawn from a finite population: the measures completed within the analysis period with associated energy impacts in the program-tracking database. This calculation uses the Finite Population Correction (FPC) factor. This factor is a reduction to the calculated variance that accounts for the fact that a relatively large fraction of the population of interest has been observed directly and is not subject to uncertainty. It is appropriate to apply precision statistics, such as confidence intervals, based on the standard error calculated in this manner when quantifying the results of the program during the study period only. The FPC factor reduces the calculated sampling error around the estimate more for smaller populations than for large.

The second calculation treats the population of interest as essentially infinite. Thus, the measures completed to date and the sample selected from them is regarded as random instances of a virtually infinite number of measures that could have been completed under the program. In this case, the FPC is not included. It is appropriate to apply standard errors calculated in this manner when applying the verification factors developed from this study to tracked savings from other years to estimate verified savings in those years.

The Gross RR, $R_{V,}$ is calculated by chaining together the installation rate and the calculation adjustment:

$$R_V = R_I R_E = \left[\frac{\sum_{j \varepsilon V} G_{Ij} w_{Vj}}{\sum_{j \varepsilon V} G_{Tj} w_{Vj}}\right]\left[\frac{\sum_{j \varepsilon V} G_{Cj} w_{Vj}}{\sum_{j \varepsilon V} G_{Ij} w_{Vj}}\right]$$

This is an example of a chained ratio estimator using a nested sample. The standard error for the chained ratio is approximated by the formula:

$$SE(AB) \cong AB\sqrt{\left[\left(\frac{SE(A)}{A}\right)^2 + \left(\frac{SE(B)}{B}\right)^2\right]}$$

(This formula overstates the standard error, because it ignores the correlation between the numerator of R_I and the denominator of R_E, which reduces the variance of the product.)

[43] For the net-to-gross ratio, the verified gross savings for measures in the FR only sample (G_{Vj}) were estimated based on the gross RRs found for measures of the same measure type in the CPSV sample.

Likewise, the Net RR, R_N, is calculated by chaining together the gross realization rate and the net-to-gross ratio:

$$R_N = R_V R_A$$

The same standard error approximation formula allows (an over-estimate of) the standard errors of each of the realization rates to be calculated from the two separate standard errors.

Ratio Estimation Example

This section provides an example of the ratio estimation procedure. The results in this section are for explanatory purposes only.

The installed savings, and engineering verified savings, are calculated at the measure level and summed to the Measure Type level for each customer in the sample that completed a survey. Attribution is collected at the measure type level and is a function of the verified measure type savings for the customer. The sample weights are applied to the measure type level savings which is the unit of analysis. Table 53 shows the reported, installed and verified savings and NTG for Example Customer A's four measures reported in the program tracking database.

Table 53: Example Customer A in CPSV and NTG Sample

Measures	Measure Type	Reported m3	Installed m3	Verified m3	NTG
Space Heat Boiler 1	Space Heat	80,000	80,000	100,000	100%
Space Heat Boiler 2	Space Heat	56,000	56,000	55,000	
Process Heat	Process Heat	150,000	150,000	120,000	80%
Steam Trap Repair	Maintenance	12,000	12,000	14,000	20%

DNV GL engineers confirmed the customer installed all of the measures that were reported by the program; therefore installed savings are equal to the reported savings. If a measure was initially reported as not installed, a second DNV GL engineer would contact the customer to verify this result. The engineering review produced adjustments to the installed savings for the first three of Customer A's reported measures, resulting in differences between the verified gross savings and installed savings for those measures.

The attribution rate is calculated for each measure type using the customer and supplier survey, if applicable, for Example Customer A using the methods that will be provided with the survey instruments. The measure type level attribution rates are then applied to the aggregated measure type level verified gross savings to estimate measure level net savings. Example Customer A received 100 percent attribution for the two space heat measures, 80 percent attribution for the process heat measure, and 20 percent attribution for the maintenance measure. Table 54 shows the verified gross and net savings for Example Customer A.

Table 54: Example Customer A Net Savings

Measure Type	Verified m3	NTG	Net m3
Space Heat	155,000	100%	155,000
Process Heat	120,000	80%	96,000
Maintenance	14,000	20%	2,800

Similar estimates are created for each customer in the sample. For this example we assume Example Customers A to F comprise the Industrial Sector sample. Table 55 shows the un-weighted customer and commercial sector savings results.

Table 55: Example Industrial Sector Measure Type Level Sample

Customer	Measure Type	Reported m3	Installed m3	Verified m3	Net m3
A	Space Heat	136,000	136,000	155,000	155,000
A	Process Heat	150,000	150,000	120,000	96,000
A	Maintenance	12,000	12,000	14,000	2,800
B	Process Heat	250,000	250,000	180,000	180,000
B	Maintenance	20,000	20,000	14,000	0
C	Space Heat	150,000	150,000	140,000	35,000
D	Process Heat	80,000	80,000	81,000	81,000
E	Space Heat	70,000	70,000	70,000	0
F	Space Heat	14,000	14,000	13,000	0

Each customer in the sample frame is assigned to a sampling stratum as described in the sampling plan. Each customer in the sample is assigned a sampling weight based on the sample design and the number of completed sample points in each stratum. Assume that Example Customers A and C each have a space heat measure in a stratum that has four measures in the sample frame. The sampling weight for the space heat measures for Customers A and C is equal to the number of customers in the sample frame stratum divided by the number of stratum customers in the sample, or 4/2 = 2. The weighted savings for each customer is equal to the weight times the savings value. Table 4 shows the weights and savings (un-weighted and weighted) for each customer in the Example Industrial Sector if we assume the measure type weights shown.

Table 56: Example Industrial Sector Measure Type Level Weighted Savings

			Reported m3		Installed m3		Verified m3		Net m3	
Customer	Measure Type	Weight	unweighted	weighted	unweighted	weighted	unweighted	weighted	unweighted	weighted
A	Space Heat	2	136,000	272,000	136,000	272,000	155,000	310,000	155,000	310,000
A	Process Heat	3.5	150,000	525,000	150,000	525,000	120,000	420,000	96,000	336,000
A	Maintenance	20	12,000	240,000	12,000	240,000	14,000	280,000	2,800	56,000
B	Process Heat	1	250,000	250,000	250,000	250,000	180,000	180,000	180,000	180,000
B	Maintenance	18	20,000	360,000	20,000	360,000	14,000	252,000	0	0
C	Space Heat	2	150,000	300,000	150,000	300,000	140,000	280,000	35,000	70,000
D	Process Heat	3.5	80,000	280,000	80,000	280,000	81,000	283,500	81,000	283,500
E	Space Heat	15	70,000	1,050,000	70,000	1,050,000	70,000	1,050,000	0	0
F	Space Heat	25	14,000	350,000	14,000	350,000	13,000	325,000	0	0

The next step is to determine program overall adjustment factors. For kWh the Industrial Sector the installation rate, engineering verification factor, and attribution adjustment factor are:

3,627,000 weighted installed m^3 / 3,627,000 weighted reported m^3 = 100% installation rate

3,380,500 weighted verified gross m^3 / 3,627,000 weighted installed m^3= 93.2% eng. verification factor

1,235,500 weighted net m^3 / 3,380,500 weighted verified gross m^3 = 36.5% attribution adjustment.

The verified gross RR is the product of the installation rate and the engineering verification factor, or 100 percent times 93.2percent = 93.2 percent for this example. The net RR is the product of the verified gross RR and the attribution adjustment, or 93.2 percent times 36.5 percent = 36.5 percent for this example.

The same principle can be applied to each Measure Type to get the Measure Type level adjustment factors. With the unit of analysis remaining the same (at the measure type level), the same process can be used to produce adjustment factors for any domain that we are able to define for the whole sample.

Applying Ratios to Domains

Ration application refers to multiplying the gross RR and net RR times the program tracking savings to produce the total verified and net savings results for a program.

The general formula for total verified gross savings is:

The general formula for total net savings is:

The body of the report discusses how to calculate the population adjustment factors, which are based on a finite, fixed distribution of projects. You can also calculate for subsets, called domains. Viewing domain-level results allows for insights into program performance that can lead to program improvements. Domain-level ratios can also be used to apply ratios and calculate overall program savings totals. The ratio results will be generated for each of the domains of interest (subsets of the population that stakeholders agree are important) and overall for each of the utilities' programs.

The level at which one applies the ratios has an effect on the overall verified and net savings estimate for each program. There are two basic approaches that we take. The first is to apply the overall program ratio. This is appropriate to retrospective evaluation where the population that the applied ratio is the same as the population of study and is static.

The second is to apply the ratio at the domain level. This is appropriate for all uses and recommended for estimating savings for programs or program years that are not the same as the population of study. Another approach is to apply the ratio at the stratum level. This is really a subset of the domain application approach where the domain used is the sample strata.

We recommend applying ratios by domains in most cases in order to improve accuracy. Assuming a sufficient sample size in each domain, domain-level precisions are usually sufficient for the approach. While

90/10 relative precision is typically the threshold targeted for an overall result, precisions usually have lower threshold for domain-level application as the resulting precision of the overall result will be better than the component parts.

If one domain has an extreme adjustment, the accuracy of the overall result is improved if domain level ratios are applied to the domain level savings. Table 57 shows an example where we apply the gross RR and net RR directly and by domains. The sample weighted savings in the example closely match the population savings: one domain, process heat, is 3.2 percent different, while the other domains are each within 3 percent and overall the difference is less than 1 percent. The ratios and resulting savings are also similar, within one percent of one another. Though the results in the example are similar, the final net savings are more accurate when calculated by domains. In the example, both space heat and maintenance measures had very different attributions from process heat and each were slightly over-represented in the weighted sample savings, which resulted in lower net savings when we applied the overall ratio directly.

Table 57: Example of Ratios Applied Overall vs. by Domains

	A	B	C	D		
Measure Type	Population m3	Sample Weighted m3	Gross RR	Net RR	Verified Gross Savings (A*C)	Net Savings (A*D)
Space Heat	1,950,000	1,972,000	99.6%	19.3%	1,943,078	375,761
Process Heat	1,090,000	1,055,000	83.7%	75.8%	912,810	826,024
Maintenance	585,000	600,000	88.7%	9.3%	518,700	54,600
Overall - Ratios Applied Directly	3,625,000	3,627,000	93.2%	34.1%	3,378,636	1,234,819
Overall - Ratios Applied by Domains and Summed	3,625,000		93.1%	34.7%	3,374,589	1,256,384
Difference			0.1%	-0.6%	4,047	-21,566

Neither applying the overall ratio directly nor by domains has an inherent systemic bias, but when the differences among the domain ratios are significant, applying by domains results in improved accuracy.

The choice between how to apply the ratios does not affect whether or which domains are reported. There is a large inherent value in looking at program results by multiple domains in order to better understand where the program is doing well and what areas have room for improvement.

Criteria for selecting domains for reporting and application

DNV GL will select the domains that are reported and those that will be applied to estimate gross and net savings for the programs.

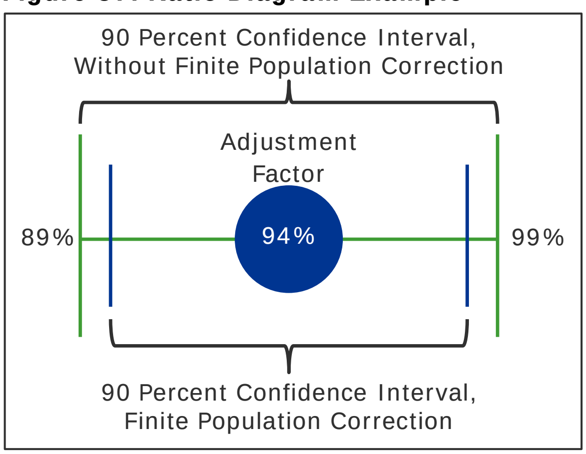
Table 58: Relevant statistics.

Term	Definition
Ratio/Adjustment factor	A point estimate of the evaluation findings expressed as a percent.
+/- or Absolute Precision	If the evaluation were repeated several times selecting samples from the same population, 90%[44] of the time the ratio would be within this range of the ratio
Confidence interval	The upper bound is defined by the ratio plus the absolute precision. the lower bound is defined by the ratio minus the absolute precision.
Relative Precision	The relative precision is calculated as the absolute precision divided by the ratio itself. By convention, relative precisions are the statistic that are targeted in sampling (ie. 90/10 is a relative precision metric)
Finite population correction (FPC)	FPC is a factor that reduces the measured error of samples drawn from small populations (less than 300). FPC applies when the ratio is applied to the same population from which the sample was drawn.

Figure 37 shows an example:

- the adjustment factor (ratio) as a blue point
- the 90 percent confidence interval *with finite population correction* (blue)
- the 90 percent confidence interval *without finite population correction* (green)

Figure 37: Ratio Diagram Example



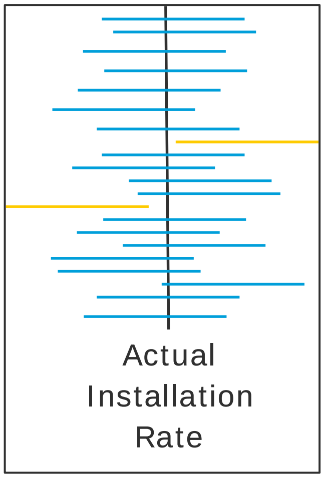
The plus/minus (±) error (%) indicated at the 90 percent confidence interval is the absolute difference between the estimated percentage and the upper or lower confidence bound. For example, in Figure 37, the ratio is 94 percent and the non-FPC 90 percent confidence interval is ± 5 percentage points (i.e., 94 percent ± 5 percent).[45] Another way of saying this is that there is a 90 percent probability that the actual ratio for

[44] 90% is the confidence limit that we are using.

[45] The critical value for calculating the confidence interval ± for each adjustment factor is determined using Student's t-distribution and n-1 for the degrees of freedom, where n is the sample size. The critical value for the gross savings adjustment factor is determined using the degrees of freedom based on the minimum sample size for the components of the adjustment factor. The gross savings adjustment factor is a product of

the next year's program lies between 89 and 99 percent. Figure 38 demonstrates this concept by showing twenty hypothetical confidence intervals calculated from twenty different samples of the same population. Eighteen out of twenty (90 percent) include the true population ratio

Figure 38. Ninety Percent Confidence Interval



Note: Each horizontal line represents a confidence interval. Yellow confidence intervals do not include the actual ratio.

The relative precision of the ratio is calculated as 5%/94% =5.3%.

For low ratios, relative precisions may be quite high, even when the confidence interval around the ratio is quite narrow. Consider a ratio of 40% with the same 5% absolute precision as in the above example. While the absolute precisions are the same, the latter ratio (40%) has a relative precision of 5%/40% =12.5%.

Because relative precisions can over-represent error for low ratios (and under-represent errors for ratios above 100%), we prefer to set thresholds for reporting and application based on the absolute precision rather than the relative precision.

For determining which ratios to report and apply we will use the following rules:

- The minimum sample size for a reporting or application domain will be five.
- The absolute precision threshold for reporting ratio for a domain will be +/- 20% at 90% confidence with FPC-on.
- The absolute precision threshold for applying ratio for a domain will be +/- 15% at 90% confidence with FPC-on for retrospective application.
- The absolute precision threshold for applying ratio for a domain will be +/- 20% at 90% confidence with FPC-off for prospective application.

---

the installation rate and the engineering verification factor. For 2-tailed estimates (ratios that could be above or below 100%) the appropriate t-stat used to calculate precision from the standard error is close to 1.645.

Reporting domains will be defined as combinations of the following categorizations where sample sizes and precisions allow:

- Stratification segments
- NTG Category (for FR and SO)
- CPSV Category (for Gross results)
- Measure types (shown in Figure 4)

Table 59 and Table 61 present the maximum number of reporting domains for the NTG results and Table 62 and Table 63 present the maximum number of reporting domains for the CPSV results.

There will be cases where some of the groups defined by a categorization have sufficient precision, while others do not. In these cases, we will combine the groups that do not meet reporting thresholds into an "other" group. For example, we may have sufficient precision to report separate ratios for Enbridge Commercial Controls, Heat Recovery and HVAC, but not enough to report the ratios for the other six measure types. In this case, we will report the three groups that we have sufficient precision for and group the rest into a "Balance of commercial/Other" group. Table 60 provides an example of how the Enbridge NTG domains presented in Table 59 could potentially be collapsed during ratio estimation.

No results will be reported that blend Union and Enbridge samples. Large Volume and RunitRight will also not be combined with other programs segments due to their different designs.

For application of CPSV results our initial list of application domains will be within stratification segment with separate domains for each CPSV category and measure type (as shown in the tables below). Those domains that meet the pre-defined precision and sample size criteria, described above, will have results applied at this level. For the rest of the list we will combine domains in the most logical manner appropriate to the ratio in order to achieve combinations that meet criteria and where possible are a meaningful grouping of measures. For example, we will combine CPSV categories within measure types and combine measure types within CPSV categories as is most reasonable given the estimation approaches used (ie if there is little difference in simple vs complex measures in the calculation method for building shell measures we would combine the simple and complex building shell first rather than simple building shell into a "simple-other" domain).

For application of NTG results the same process will be used as for CPSV, but with the NTG category substituting for the CPSV category.

Table 59. Enbridge NTG Domains

Utility	Program	NTG Category	Measure Type
Enbridge	Commercial	Action	Operational Improvements
Enbridge	Commercial	Equipment	Building Shell
Enbridge	Commercial	Equipment	Controls
Enbridge	Commercial	Equipment	HVAC
Enbridge	Commercial	Equipment	Heat Recovery
Enbridge	Commercial	Equipment	Steam and Hot Water
Enbridge	Industrial	Action	Operational Improvements
Enbridge	Industrial	Equipment	Building Shell
Enbridge	Industrial	Equipment	Controls
Enbridge	Industrial	Equipment	Greenhouse
Enbridge	Industrial	Equipment	HVAC
Enbridge	Industrial	Equipment	Heat Recovery
Enbridge	Industrial	Equipment	Other Equipment
Enbridge	Industrial	Equipment	Process Heat
Enbridge	Industrial	Equipment	Steam and Hot Water
Enbridge	Multi-Residential	All	Controls
Enbridge	Multi-Residential	All	HVAC
Enbridge	Multi-Residential	All	Heat Recovery
Enbridge	Multi-Residential	All	Operational Improvements
Enbridge	Multi-Residential	All	Steam and Hot Water
Enbridge	RunitRight	Action	RunitRight

Table 60. Example of Potential Enbridge NTG Domain Collapsing

Utility	Program	NTG Category	Measure Type
Enbridge	Commercial & Multi-Residential	Equipment	Controls
Enbridge	Commercial & Multi-Residential	Equipment	HVAC
Enbridge	Commercial & Multi-Residential	Equipment	Steam and Hot Water
Enbridge	Commercial	Equipment	Other Commercial Equipment
Enbridge	Industrial	Action	Operational Improvements
Enbridge	Industrial	Equipment	Controls
Enbridge	Industrial	Equipment	Heat Recovery
Enbridge	Industrial	Equipment	Other Industrial Equipment
Enbridge	Multi-Residential	All	Multi-Residential Other
Enbridge	RunitRight	Action	RunitRight

Table 61. Union NTG Domains

Utility	Program	NTG Category	Measure Type
Union	Commercial	Action	Controls
Union	Commercial	Action	Maintenance
Union	Commercial	Action	Optimization
Union	Commercial	Action	Steam and Hot Water
Union	Commercial	Equipment	Building Shell
Union	Commercial	Equipment	Controls
Union	Commercial	Equipment	HVAC
Union	Commercial	Equipment	Heat Recovery
Union	Commercial	Equipment	New Construction
Union	Commercial	Equipment	Other Equipment
Union	Commercial	Equipment	Steam and Hot Water
Union	Industrial	Action	Controls
Union	Industrial	Action	HVAC
Union	Industrial	Action	Maintenance
Union	Industrial	Action	Optimization
Union	Industrial	Action	Steam and Hot Water
Union	Industrial	Equipment	Ag and Greenhouse
Union	Industrial	Equipment	Building Shell
Union	Industrial	Equipment	Controls
Union	Industrial	Equipment	HVAC
Union	Industrial	Equipment	Heat Recovery
Union	Industrial	Equipment	Other Equipment
Union	Industrial	Equipment	Process Heat
Union	Industrial	Equipment	Steam and Hot Water
Union	Large Volume	Action	HVAC
Union	Large Volume	Action	Heat Recovery
Union	Large Volume	Action	Maintenance
Union	Large Volume	Action	Optimization
Union	Large Volume	Action	Other Equipment
Union	Large Volume	Equipment	Ag and Greenhouse
Union	Large Volume	Equipment	Building Shell
Union	Large Volume	Equipment	Controls
Union	Large Volume	Equipment	HVAC
Union	Large Volume	Equipment	Heat Recovery
Union	Large Volume	Equipment	New Construction
Union	Large Volume	Equipment	Other Equipment
Union	Large Volume	Equipment	Steam and Hot Water
Union	Multi-Family	All	Controls
Union	Multi-Family	All	New Construction
Union	Multi-Family	All	Steam and Hot Water

Table 62. Enbridge CPSV Domains

Utility	Program	CPSV Category	Measure Type
Enbridge	Commercial	Complex	Building Shell
Enbridge	Commercial	Complex	Controls
Enbridge	Commercial	Complex	HVAC
Enbridge	Commercial	Complex	Heat Recovery
Enbridge	Commercial	Complex	Operational Improvements
Enbridge	Commercial	Complex	Steam and Hot Water
Enbridge	Commercial	Simple	Building Shell
Enbridge	Commercial	Simple	Controls
Enbridge	Commercial	Simple	HVAC
Enbridge	Commercial	Simple	Heat Recovery
Enbridge	Commercial	Simple	Operational Improvements
Enbridge	Commercial	Simple	Steam and Hot Water
Enbridge	Industrial	Complex	Building Shell
Enbridge	Industrial	Complex	Controls
Enbridge	Industrial	Complex	Greenhouse
Enbridge	Industrial	Complex	HVAC
Enbridge	Industrial	Complex	Heat Recovery
Enbridge	Industrial	Complex	Operational Improvements
Enbridge	Industrial	Complex	Other Equipment
Enbridge	Industrial	Complex	Steam and Hot Water
Enbridge	Industrial	Simple	Building Shell
Enbridge	Industrial	Simple	Greenhouse
Enbridge	Industrial	Simple	HVAC
Enbridge	Industrial	Simple	Heat Recovery
Enbridge	Industrial	Simple	Operational Improvements
Enbridge	Industrial	Simple	Other Equipment
Enbridge	Industrial	Simple	Process Heat
Enbridge	Industrial	Simple	Steam and Hot Water
Enbridge	Multi-Residential	Complex	Controls
Enbridge	Multi-Residential	Complex	HVAC
Enbridge	Multi-Residential	Complex	Heat Recovery
Enbridge	Multi-Residential	Complex	Steam and Hot Water
Enbridge	Multi-Residential	Simple	Controls
Enbridge	Multi-Residential	Simple	HVAC
Enbridge	Multi-Residential	Simple	Heat Recovery
Enbridge	Multi-Residential	Simple	Operational Improvements
Enbridge	Multi-Residential	Simple	Steam and Hot Water

Table 63. Union CPSV Domains

Utility	Program	CPSV Category	Measure Type
Union	Commercial	Complex	Building Shell
Union	Commercial	Complex	Controls
Union	Commercial	Complex	HVAC
Union	Commercial	Complex	Heat Recovery
Union	Commercial	Complex	Maintenance
Union	Commercial	Complex	New Construction
Union	Commercial	Complex	Optimization
Union	Commercial	Complex	Other Equipment
Union	Commercial	Complex	Steam and Hot Water
Union	Commercial	Simple	Building Shell
Union	Commercial	Simple	Controls
Union	Commercial	Simple	HVAC
Union	Commercial	Simple	Heat Recovery
Union	Commercial	Simple	Maintenance
Union	Commercial	Simple	Optimization
Union	Commercial	Simple	Other Equipment
Union	Commercial	Simple	Steam and Hot Water
Union	Industrial	Complex	Ag and Greenhouse
Union	Industrial	Complex	Building Shell
Union	Industrial	Complex	Controls
Union	Industrial	Complex	HVAC
Union	Industrial	Complex	Heat Recovery
Union	Industrial	Complex	Maintenance
Union	Industrial	Complex	Optimization
Union	Industrial	Complex	Other Equipment
Union	Industrial	Complex	Process Heat
Union	Industrial	Complex	Steam and Hot Water
Union	Industrial	Simple	Ag and Greenhouse
Union	Industrial	Simple	Building Shell
Union	Industrial	Simple	Controls
Union	Industrial	Simple	HVAC
Union	Industrial	Simple	Heat Recovery
Union	Industrial	Simple	Maintenance
Union	Industrial	Simple	Optimization
Union	Industrial	Simple	Process Heat
Union	Industrial	Simple	Steam and Hot Water
Union	Large Volume	Complex	Ag and Greenhouse
Union	Large Volume	Complex	Building Shell
Union	Large Volume	Complex	Controls
Union	Large Volume	Complex	HVAC
Union	Large Volume	Complex	Heat Recovery
Union	Large Volume	Complex	Maintenance
Union	Large Volume	Complex	New Construction
Union	Large Volume	Complex	Optimization
Union	Large Volume	Complex	Other Equipment
Union	Large Volume	Complex	Steam and Hot Water
Union	Multi-family	Complex	Controls
Union	Multi-family	Complex	Heat Recovery
Union	Multi-family	Complex	New Construction
Union	Multi-family	Complex	Other Equipment
Union	Multi-family	Simple	Building Shell
Union	Multi-family	Simple	Controls
Union	Multi-family	Simple	HVAC
Union	Multi-family	Simple	Steam and Hot Water

APPENDIX F DATA COLLECTION INSTRUMENT OUTLINES

This appendix provides outlines of the topics to be included in each of the data collection instruments.

Program Energy Advisors

The outline for the initial Program Energy Advisor interview is presented below:

- **Introduction**
- **General Interactions**
  - Their typical responsibilities
  - The nature of their routine communications and interactions with Custom C&I customers and how this might vary with the size of the customer or the customer type (e.g. chain stores)
- **DSM Program Promotion**
  - How they target program recruitment at Custom C&I customers of certain types or in certain areas
  - Nature of program recruitment; communication type by customer size and rate class
  - How they become aware of potential energy efficiency projects
  - How they promote energy efficiency
  - How they identify which customer representatives are the key project decision makers
  - Which customer representatives are the key project decision makers and how this might vary depending on company/organization size or type
  - At what stage in project development they typically get involved with a project
  - How many projects are typically rejected
  - What are the barriers to program participation and how they try to mitigate them
  - What information, financial incentives or technical assistance they offer to Custom C&I customers for energy efficiency projects
  - What they perceive to be their most valuable contributions to the development of energy efficiency projects
  - How frequently they rely on program technical support staff for project support
  - How closely they monitor the progress of active projects
    - If there is any evidence of project cancellations due to free ridership
  - Whether they have received any training or guidance on how to minimize free ridership
  - Whether there are any warning signs that a project might be a free rider
  - What roles vendors play in project identification
    - How Energy Advisors interact with vendors
    - How vendors work to mitigate free ridership
- **Closing**

Participating Customers

The outline for our Participant IDI is presented below, with details to be provided in the survey instruments and methods memos.

OUTLINE:

- **Introduction**
- **Informed Respondent**
  - Cite specific project, determine involvement

- Program awareness
- Equipment choice
- Role
- Responsibilities
- Ask about how long at company (since before the project date?)
- Identify names of other primary project contacts, for potential follow up conversation with DNV GL engineer

- **Organization Policies and Purchase Procedures**
- **General Program Awareness and Interactions**
- **Custom Program Savings Verification**
  - Specific questions determined based on engineer review of project documentation
- **Specific project verification (Framing)**
  - When first considered?
  - Reasons for project?
  - Major sources of info?
  - The general decision-making concerning energy related purchases and practices
  - Who in their company or organization makes decisions about equipment replacement and retrofit projects and how this might vary with the size or cost of the project
  - What information sources are used in making these decisions
  - Whether the company/organization has any formal requirements or informal guidelines about the purchasing of energy using equipment and, if so, what are these requirements/guidelines
  - Whether their company has a corporate “green “ mandate
  - The development of the specific program-incentivized project
  - Where the idea for the project originated and who were the key persons involved in the project conception -- whether within the participant’s company/organization or without (e.g., vendors, Custom C&I program Energy Solutions Consultants)
  - Who was involved in the planning and development of the project details
  - Who was involved in the decision to go ahead with the project
  - At what stage in this project conception, planning and development process did the Custom C&I program get involved
  - Whether the program provided any services to the respondent’s company/organization beyond the financial incentives (e.g., training, audits, technical assistance, helping find a vendor, selling the project to upper management, etc.). To inform the free ridership questions, the interviewers will have information on all program activities reported by the tracking databases, but this query is designed to collect information on program activities that may be unreported and also to find out which program activities were top-of-mind for the respondent.
  - Whether the project changed from its original conception and what these changes were and why they were made

- **Direct attribution battery**
  - Determining the overall influence of the program, along with program effects on
    - Timing

        - Efficiency
        - Sizing or Quantity
    - Long term program effect (secondary attribution)
- **Spillover battery (for customers in both FR and Spillover samples)**
    - Inquire about additional projects *after* other projects[46]
        - First check to ensure not incentivized
        - Project type
        - Project data
        - Project contact
        - Project location
        - Project dates
- **Firmographics**
    - Business type and
    - Business size (ft2 and # of employees)
    - whether they lease or own their facilities
- **Closing**

Participating Vendors

IDIs with up to 80 of these vendors will follow the following outline:

- **Introduction**
- **Informed Respondent**
- **Company background**
    - Which products or services they sell
    - Which types of C&I customers they typically do business with
    - What the size of their company is
- **Sales and marketing**
    - How they promote energy efficiency
    - How they identify which customer representatives are the key project decision makers
    - Which customer representatives are the key project decision makers and how this might vary depending on company/organization size or type
    - What role the Custom C&I program incentives play in their sales pitches
- **General program involvement and influence (General Framing)**
    - How they became involved with the Custom C&I program

[46] Information collected regarding additional projects will be used not only to calculate spillover, but to check against program records and ensure that the project was not a tracked project with direct attribution.

- Why they became involved with the program
- How frequently they offer program incentives
- How frequently they interact with program staff
- How they keep track of Custom C&I program incentives and requirements
- Whether the Custom C&I programs have provided them with any sales leads
- Whether they have received any training from the program
- Whether there are other services that the program provide them
- To what degree the Custom C&I program incentives and other services influence the implementation of energy efficiency projects in the C&I sector
- What types of C&I customers are more likely to be influenced by program incentives and which types are less likely to be influenced
- Whether they are offering energy efficient products or services through the program that they did not offer before becoming involved with the program
- Whether they are recommending energy efficient products or services more frequently now than they did before becoming involved with the program
- Whether they have suggestions as to what kind of interventions would increase the program's influence in the project

- **Project Specific Framing**
  - Whether they were involved in originating the project idea and, if so, what was their role (informed respondent)
  - Whether they were involved in planning and the development of the project details, and if so, what was their role
  - Whether they were involved in the decision to go ahead with the project, and if so, what was their role
  - At what stage in this project conception, planning and development process they got involved
- **Project Specific Attribution**
  - Whether the availability of the Custom C&I program financial incentives or other Custom C&I program services (e.g., training, audits, technical assistance, helping find a vendor, selling the project to upper management, etc.) that the participant received had any influence on
    - timing or
    - efficiency
    - quantity/size of the project

    and if so, what was the nature of this influence.
- **Closing**

APPENDIX G CPSV SITE REPORT TEMPLATES

The embedded documents below are the draft templates for CPSV that will be used for the TSER and Onsites respectively.

CPSV Onsite Site Report Template - OE

CPSV TSER Site Report Template - OE

APPENDIX H CPSV RIGOUR LEVELS

The CPSV plan calls for two types of data collection: telephone-supported engineering review (TSER) and onsite. There are adjustments that might entail more or less work at each site. Table H details likely engineering effort levels for the standard, increased, and decreased levels. The levels of effort are averages. Some sites may require substantially more effort, while some sites may entail less effort.

Based on the tracking data we have identified the simplest projects as a level of stratification and will used TSER interviews to verify the projects at these sites. The more complex onsite sample will also have varying degrees of effort requirements in order to allow more effort at more complex sites.

Table H: M&V Description for Proposed Engineering Effort Levels

Effort Level	Description
Telephone-supported engineering review (TSER)	Lower rigour projects. Application desk review, telephone interviews, possible revised engineering calculations; primarily for qualitative assessment.
Standard Onsite	Simpler projects. Detailed application review, on-site verification, collection of data on key parameters, revised engineering calculations, billing data analysis, and possible spot measurements.
Higher Rigour Onsite	Small, medium and large scale projects that may or may not require monitoring or metering. Detailed application review, on-site verification, collection of data on key parameters, revised engineering calculations, billing data analysis, and possible spot measurements / short term post monitoring.
Very High Rigour Onsite	Largest and most complex projects. Detailed application review, on-site verification, collection of data on key parameters, billing/interval data analysis, calibrated simulation models, spot measurements, long-term post monitoring, pre-verification and short-term measurement. May require larger teams, including senior staff and multiple site visits.

Most site-specific impact evaluation efforts for Standard Onsite points will fall into the category of lower rigor level of effort. However, there are exceptions and adjustments that might entail more or less work at each site. During the file review adjustments of this sort should be noted and the sites will be reviewed by the engineering team lead (Phani Pagadala) to determine which level of rigour is required. Up to 20 sites (primarily Large Volume) will receive higher rigour onsites and up to two sites will receive very high rigour onsites to establish the relative value of increased rigour levels in future evaluation.

Each site will be assigned a single point of contact (POC) for the purposes of communications with the customer, the utility and within CPSV itself. The POC will be a more senior engineering team member who is experienced in the energy efficiency field (preferably a registered professional engineer) and will be responsible for co-ordinating the work of their team, tracking progress on each project review, becoming intimately familiar with the documentation and technical requirements of the work to be performed, ensuring that quality control procedures are implemented, and reporting on project review progress and any issues to the engineering team lead (Phani Pagadala).

Higher rigour sites could involve the addition of elements such as:

- A fully specified regression analysis of consumption information from utility bills with inclusion/adjustment for changes and background variables over the time period of the analysis that could potentially be correlated with the gross energy savings being measured.
- Twelve (12) months post-retrofit consumption data are required.
- Twelve (12) months pre-retrofit consumption data are required, unless program design does not allow pre-retrofit billing data, such as in new construction. In these cases, well-matched control groups and post-retrofit consumption analysis is allowable.
- Sampling must be adequate (in general, a minimum of six data points will be required) for a valid regression-based estimate.
- Building energy simulation models that are calibrated as described in IPMVP Option D requirements. If appropriate, evaluators may alternatively use an engineering model with calibration.
- Retrofit isolation engineering models as described in IPMVP Option B requirements.

APPENDIX I TEC METHODOLOGY DECISIONS

This appendix contains the discussion history of methodological decisions that were left unresolved following the initial project kickoff meeting in 2014.

Date: June 9, 2015

Memo to:
Bob Wirtshafter, Independent Member
Ted Kesik, Independent Member
Chris Neme, Green Energy Coalition
Jay Shepherd, School Energy Coalition
Julie Girvan, Consumers Council of Canada
Ravi Sigurdson, Enbridge Gas
Marc Hull-Jacquin, Enbridge Gas
Tina Nicholson, Union Gas
Meredith Lamb, Union Gas

Copy:
Mimi Goldberg, DNV GL
Tammy Kuiken, DNV GL

Prep. by: Ben Jones, DNV GL

Ontario Gas NTG Evaluation Kickoff Meeting Items

This memo memorializes the discussions of unresolved parking lot items from the 2014 Ontario Gas Net-to-Gross Evaluation kickoff meeting. It is intended to identify which of the items were resolved, assigned (action items), or discussed, but ultimately re-tabled at the meeting (parking lot items). The initial Parking Lot items, DNV GL and TEC takes are retained for context (in grey).

Parking Lot Items Discussed

Several Items at the kickoff meeting were discussed but ultimately tabled without a resolution. The "DNV GL Take" below has the evaluation team's initial thoughts about how each issue should be addressed, while the TEC Take provides the TEC's consensus prior to the follow up meeting.

1. *How much contact should the evaluation have with program staff regarding specific projects?*
    - The utilities would like the evaluation to meet with program staff to discuss the specifics of all projects, not just the specific large or complicated ones or the ones we deem need additional information.
    - Other TEC members worry that too much contact with the utility reps will lead to a biased evaluation.
    - *DNV GL Take* – For complex projects, understanding the timing and specifics of the program's interactions with the customer provides the evaluation with the ability to tailor questions prior to the core attribution sequence to the specific customer experience. These custom questions will be phrased to remind the customer about the interactions, while being careful not to bias the customer's responses to the core attribution questions.
        - The TEC is not in a position to provide endorsement on this point at this time. Discussion with DNV is required.
    - TEC Take:
        - DNV should determine the extent of contact it requires with utility program staff, in order to be fully informed on the customer's relationship with each utility prior to

conducting the Net to Gross survey, given the complexity of the project and the contents of the project files. DNV will follow up as required with the utilities.
    - Discussion with DNV is required on the highlighted issue above regarding tailoring questions to remind customers about their interactions with the utility.
- June 11 Follow up discussion results (not discussed on July 16):
    - **TEC Action Item:** TEC to discuss guidelines for framing[47] questions and usage of information from program in probes. Decision required prior to survey instrument development.
    - Open questions on usage of
        - framing questions to remind customer of decision making process
        - program-supplied information in framing question-related probes
    - Rationale for questions and scoring to be provided with survey instruments once drafted.
    - Notes for consideration:
        - Specific questions and probe instructions will be reviewed by TEC prior to fielding interviews. Initial decision for TEC is whether or not the general approach is acceptable.
        - Important to remember that this section of the interview is not part of the scoring algorithm. It is intended to help respondents recall a project and process that may have occurred a few years ago. It is not intended to push the participant into giving more credit to the program than they would if we asked the NTG questions when the decision was fresh in mind. Aiding participant recall through framing questions attempts to remove an aspect of self-report surveys that can potentially bias results against giving programs credit for the decision to install EE equipment.

2. *Can the evaluation determine which portions of the attribution were due to financial incentives, which were other services, etc?*
    - *DNV GL Take* – The proposed attribution approach and the current scope of the evaluation does not allow for proportionally assigning attribution credit to different program influences. Attribution is also not a zero sum game: both technical and financial assistance may be necessary for a project to proceed; the absence of either one could be enough to prevent a customer from going forward. The surveys and interviews will gather qualitative information about the influence of different program activities on projects. The evaluation will report these (anonymized) responses relative to the final attribution scores in order to provide the TEC and programs some feedback in this area.
    - TEC Take:
        - The TEC would like DNV GL to gather and report on qualitative information about the influence of different program activities on projects to the extent that can be done within the defined project scope and budget.
        - The related item of the construct of the raw data and whether or not it will be shared requires discussion with DNV.
    - June 11 Follow up discussion results (not discussed on July 16):

[47] Framing questions are those that remind the customer of the decision-making process and are not used in the attribution scoring.

  - TEC Take #1 resolved
  - TEC Take #2 tabled
  - Notes on TEC Take #1 decision:
    - DNV GL to include qualitative discussion in text of participant-reported reasons for results describing NTG and spillover analysis results. Qualitative information will be provided with context such as number of respondents who provided a given reason.
    - Some open-ended responses will be scrubbed and provided in report to add context and support to the results.
- Potential TEC or utility interest in later additional analysis using the data collected
  - Notes on TEC Take #2 discussion:
    - Data must be anonymized before delivery to TEC
    - Decision to be based on usefulness and cost

3. *Do we want to make a concerted effort to talk to self-direct customers who only spent a portion of their incentive money? As opposed to customers who used it all because they lose it otherwise.*
   - *DNV GL Take* – Assuming that data on this topic is available to the evaluation and categorizing customers by proportion of incentive money spent is straightforward, the evaluation could potentially stratify based on this metric, or not stratify based on it but still attempt to report results for each group separately. Stratification by a categorization allows the evaluation to ensure that one group is not over-represented in the final weighted results, given the potential that there are meaningful NTG differences based on this categorization it likely will make sense to stratify by it if possible.
   - <u>TEC Take</u>:
     - The TEC agrees that the sample for Union's self-direct customers should be representative of the entire self-direct program, including both self-direct customers who spent all their allocated funds and those who spent only a portion of them. The TEC will defer to DNV's expert judgment regarding whether stratification based on this variable is appropriate to maximize the accuracy of a NTG for the entire program.
     - DNV should also note that a portion of Union's self-direct funds were not used by customers to which they were initially allocated. Those unused funds were then dispersed via an aggregated pool approach where projects were supported based on their lifetime natural gas savings and cost effectiveness. Again, the TEC will defer to DNV's expertise regarding how to best incorporate NTG impacts from the aggregate pool approach into an NTG for the entire program.
   - June 11 Follow up discussion results: Resolved
     - DNV GL to use expert judgment in making decision
     - Final stratification to be representative
     - The aggregated pool approach may change participant behavior in terms of what measures they do when
4. *The utilities report lifetime savings; should the evaluation use a dual baseline net-to-gross calculation? If so, how will the evaluation determine existing efficiency baseline savings without doing the full verified gross savings calculation process?*
   - *DNV GL Take* – If the program tracks dual baseline savings, the evaluation could use the information in our net-to-gross calculations. Otherwise, we might be able to use another

approach, such as assuming a ratio of the difference in savings from the dual baselines based on another study.

- TEC Take:
    1. The 2008 Summit Blue Free Ridership Study accounted for advancement through the concept of partial free ridership. Thus, the utilities do adjust savings for advancement but do not take a dual baseline approach. The TEC would like to discuss with DNV the alternative approaches.
    2. This item requires discussion with DNV to determine the implications of this for the NTG study. Note that the utilities are about to face a new DSM Framework in 2015. We do not yet know how goals will be set in that framework. It is possible that they will be set differently than the current lifetime savings (CCM) approach. Thus, we would like to know if it would be possible to adjust an NTG result computed for a CCM metric to a TRC metric if such a change was necessitated by a change in the DSM framework? Is DNV able to do a lifetime and annual calculation?
- June 11 Follow up discussion results: Not Resolved
    - **DNV GL Action Item:** DNV GL to provide simplified explanation of the two approaches and the pros and cons of each.
    - **TEC Action Item:** TEC to decide whether to pursue both methods, or select one. Resolution needed prior to starting analysis.
    - Providing both LCNS and Y1NS results is relatively straight forward, however using LCNS for these programs would require a general rather than specific estimation approach for dual baselines, making it less accurate than its original intended design
- July 16 Follow up discussion results: Resolved
    - Study will use Y1NS method with lifetime savings

5. *There is dissention about when influence occurred and what it means for NTG, largely around projects that receive incentives and are free riders in the current program year but were not free riders when they participated the first time in a past program year. How many historical program years should be taken into account by the study in determining NTG?*
    - *DNV GL Take* – This is a crucial question for the evaluation. What type of NTG are we measuring? If the study is intended to capture current program effects then a short time horizon should be used in framing questions for customers and vendors. If the study is intended to show the cumulative effect of the programs over time, then a longer time horizon should be used and past program participation and the effect of that participation on recent within-program projects should be taken into account. The surveys can be designed to capture either type of NTG, but we do not recommend attempting to capture both the current program and cumulative program versions of attribution and spillover at once: this would result in longer, more confusing surveys for customers. This is a critical item to resolve prior to developing survey instruments and interview guides. The decision as to which NTG type to pursue is ultimately a policy decision that may come down to the intent of the Ontario Board of Energy's definition of Net-to-Gross.
    - TEC Take: The TEC is not in agreement on what type of NTG the study is measuring (cumulative program effects vs. current program effects). In the absence of both TEC consensus and direction from the Ontario Energy Board, would it be possible in the current

budget and scope to calculate the NTG both ways capturing both current and cumulative effects? During discussions, the TEC considered the issues of:

  - Long life cycle projects versus projects of a repetitive nature;
  - The continuous improvement focus of the custom program design;
  - Asymmetrical treatment of accounting for utility influence and savings using a short term approach; and
  - Projects in which the lifetime claim accounts for all behaviours and years versus those projects that do not.

- June 11 follow up discussion results: Not Resolved
  - **TEC Action Item:** Decide which approach is preferred or whether surveys and interviews should attempt to capture both types of program effects. Decision required prior to survey instrument development.
  - Specific program activities that influenced the project we're looking at in this program year are taken into account no matter when they had influence. This applies primarily to the long life cycle projects.
  - Both types of program effects are important. Capturing both is interesting and also allows flexibility if OEB later decides in favor of one approach over the other.
  - Potentially could capture both types for specific projects or project types where the difference is likely to be greatest (recurring O&M for instance)
  - Deciding on one or the other prior to reporting is important to avoid higher stakes debates once results are known
- July 17 Follow up discussion results: Partially Resolved
  - TEC approves capturing long sales cycle program effects in estimation of free ridership
  - **TEC Action Item:** Continue discussion of how to capture "in program" spillover: projects incentivized in current year that were free riders based on current year program effects, but attributable to prior program participation. Consensus appeared to be that the study should capture these effects as an incremental portion of net savings so that it can be removed if need be. How to label these savings is also unresolved.
- June 9 2015 Subcommittee meeting results: Resolved (pending broader TEC approval)
  - Subcommittee recommends that the primary objective of the free ridership estimation will be to capture the effect of the program(s) on the current project. The effect on the current project of prior and indirect program experience will be captured in a secondary, less rigorous question sequence. The work plan will propose specifics for operizationalizing this approach.

6. *Should the evaluation do spillover analysis with the large industrial customers in Union Gas' new self-direct program, even though there hasn't been much time for them to complete projects? It would give the TEC something to use going forward, even if it's understated.*
   - *DNV GL Take* – Most of the data collection with this group of customers is likely to be via in depth interviews (rather than CATI surveys), which offers flexibility to inquire qualitatively about spillover potential for the program going forward as well as whether any spillover has already occurred as a result of the 2013 program. Another possible option is to ask these customers about spillover from previous program experiences in 2011 and 2012, and then

ask how the current program design would change the likelihood for future spillover. We recommend leaving this as an open question until the evaluation team learns more about the program and the overlap in customers in the 2011/2012 programs and the 2013 program.

- TEC Take: The TEC agrees to leave this as an open item until DNV has had a chance to learn more about Union's self-direct program. After DNV's review of the program, the TEC will expect a recommendation from DNV on how to perform the spillover analysis on Union's self-direct program.
- June 11 follow up discussion results (not discussed on July 16): Tabled.
    - **DNV GL Action Item**: DNV GL will recommend to the TEC a course of action for estimating spillover for the Union self-direct program once more information has been reviewed.

NAVIGANT

Net-to-Gross Policies:
Cross-Cutting Jurisdictional Review

Prepared for:

Enbridge Gas Distribution, Inc.

and

Union Gas Limited

Prepared by:

Daniel M. Violette, Ph.D.
Navigant Consulting, Inc.
1375 Walnut Street
Suite 200
Boulder, CO 80302
303.728.2500

And

Scott Dimetrosky and Lauren Miller Gage
Apex Analytics, LLC
2500 30th Street, Suite 207
Boulder, CO 80302
303.590.9888

December 14, 2017

TABLE OF CONTENTS

1. INTRODUCTION

This report reviews energy efficiency (EE) policies across different jurisdictions related to the estimation of net-to-gross (NTG) values and their application within an integrated policy framework. The objective is to better understand the NTG landscape and provide information that might inform Ontario policy decision-making related to NTG in evaluation, program planning, use in measuring progress toward savings targets, and in determining shareholder incentives.

Navigant, Inc. and Apex Analytics (Navigant Team) performed two analyses:

1. Case study analyses for three states: Massachusetts, California, and Illinois. These states were selected because, like Ontario, they have a long history of large-scale utility efficiency programs and have addressed many of the same issues regarding NTG policy that Ontario is facing today. In addition, each state has revised its policies in the past few years, resulting in recent experience in assessing NTG issues and their relationship to EE targets and incentives. The team interviewed experts in each state, including evaluators, utilities, and regulators, all experienced professionals directly involved in developing and applying NTG results or developing NTG policy in their respective states.
2. To place these case studies in a larger context, the team conducted an update to a 2015 review of NTG policies across the US.[1] This update uses state energy policy documents and websites, evaluation reports, and prior studies that sought to summarize NTG policies across states in the US.

This report is organized in the following sections:

- Section 2 presents the results from the in-depth analysis of the case study states of Massachusetts, California, and Illinois.
- Section 3 presents the findings from a comprehensive jurisdictional scan of NTG policies and the relationship of those policies to other demand-side management (DSM) requirements.
- Appendix A contains additional detail on the high level, state-by-state review.

[1] This research represents an update to a 2015 Navigant study -- Navigant. *Iowa Energy-Efficiency Net-to-Gross Report*. Prepared for the State of Iowa Department of Commerce Utilities Board. 2015. Link: https://efs.iowa.gov/cs/groups/external/documents/docket/mdax/mjax/~edisp/1201494.pdf

2. CASE STUDIES: MASSACHUSETTS, CALIFORNIA, AND ILLINOIS

This section describes the framework, history, and expert opinions regarding NTG policies, including shareholder incentives, timing of evaluation results, decision-making, NTG methods, and market effects from a closer examination of NTG policies in Massachusetts, California, and Illinois.

2.1 Approach

Massachusetts, California, and Illinois were selected as each has sizeable investments in EE programs, has implemented programs over a long period of time, and has a long history of program evaluation. In addition, each state has recently used stakeholder processes to revise the methods used to estimate NTG and how NTG is applied for incentive calculations and program planning. As a result, each has recently contemplated the pros and cons of different NTG uses and approaches while revising their policies and approaches.

The Navigant team examined state energy policy documents and websites, evaluation reports, and prior studies that sought to summarize the NTG policies in these three states. In addition, experts from each case study state were interviewed to expand the depth of the information and develop a better understanding of how policies were implemented in practice. Respondents included evaluators, utility experts, and regulators, all experienced professionals directly involved in developing and applying NTG results, or developing NTG policy in their respective states. Interviews were conducted in November 2017.

2.2 Summary of Findings

This case study review produced four overarching themes and four cross-cutting findings related to the application of NTG estimation methods. The four overarching themes are presented below:

Theme 1. **Applying NTG estimates for incentive and target calculation.** All three states have shareholder incentives, with California revising existing incentives programs in the past 2 years and Illinois designing an incentives program scheduled to start in 2018. Net savings and NTG are one of the factors used in determining incentives and judging progress toward targets in each state; however, each state has developed structures that reduce the influence of after-the-fact (i.e., retrospective) application of NTG estimates. Massachusetts and Illinois do not apply retrospective NTG estimates for determining shareholder incentives, and California's new incentive structure reduces the effect of retrospective application of NTG estimates by making it only one of four factors that are used to determine incentives and by using retrospective NTG only for select programs.

Theme 2. **Aligning savings estimates with ratepayer value.** Experts interviewed indicated that the policy purpose in using net savings (either prospectively or retroactively) is to align utilities' goals with ratepayer value. In this context, NTG estimates provide information used to inform EE investment decisions and program plans. However, the experts also reported that using net savings puts pressure on the accuracy of NTG evaluation results, specifically when the results are applied retrospectively to assess achievement of savings targets and calculation of utility incentives. Applying NTG estimates prospectively reduces uncertainty for utilities by eliminating the risk of a retroactive application of a different NTG ratio than that assumed in program planning and avoids the controversy and arguments over attribution issues that have occurred in other jurisdictions.

Theme 3. **Finalizing NTG estimates.** The process used to finalize NTG estimates to be used in shareholder calculations, program re-design, and EE targets involved more than taking the result of a study. In all three states, stakeholders had the opportunity to question, challenge, and suggest modifications to the initial estimates produced by an evaluation, management, and verification (EM&V) study. For example, in Massachusetts there is a debate and discussion or results with the goal of gaining a consensus value. If there is no agreement on the value, the regulator's evaluation consultant makes a determination considering issues raised by stakeholders. In addition, there are avenues to appeal this determination in Massachusetts if the decision is viewed as unreasonable by involved parties.

Each of the three states examined has a process by which agreement is sought among stakeholders as part of the process to finalize the NTG estimates. The experts interviewed indicated that these processes were driven by the view that NTG estimation methods all face challenges in application. This was particularly true for the self-report survey method, which is most often used for C&I customer programs, and where judgment is required to develop question batteries and survey protocols and translate survey responses into NTG values.

Theme 4. **Collaborative Overall stakeholder process**. In addition to transparency and review of final NTG estimates, the stakeholder processes in each state were predicated on defined, open stakeholder processes that build confidence in the NTG estimation process. All three case study states have adopted processes that develop agreed-upon approaches for estimating NTG as well as processes for finalizing NTG estimates. Pre-defined methods include agreeing on self-report questions and NTG assignment algorithms, incorporating multiple influence factors (e.g., program, trade ally, and market based), and scoring algorithms tested through sensitivity analyses. These common algorithms allow NTG results to be compared across programs and over time. Stakeholder processes in the case study states were predicated on transparency and discussion to build confidence in the NTG estimation process and final NTG values. Experts in each state report that the prospective application of results, combined with the consistency of the pre-defined methods and transparent stakeholder processes throughout all stages of evaluation, has created more certainty and confidence for stakeholders in terms of the actions needed meet EE targets. Additional benefits of the re-designed processes in these three states includes a more collaborative stakeholder effort and a focus on continuous improvement of programs compared to the processes that states had in prior years.

Four cross-cutting NTG methods findings are shown below:

1. ***Selected NTG methods.*** Each case study state uses methods other than self-report surveys, such as randomized control trials (RCTs) and comparison states, but these are typically used for residential or mass-market commercial products. All experts noted concerns with self-report methods but said that the primary method for custom project NTG is self-report survey methods due to the unique nature of commercial and industrial (C&I) custom projects.
2. ***Mitigating issues with self-report methods.*** Given concerns with self-report methods, experts noted the following approaches are used to improve the accuracy of self-report studies:
    - **Fast feedback:** Fast feedback refers to survey methods where the respondents are asked about factors influencing their participation in a program at a time near to when they participated—e.g., within 3 months of completing participation. Experts noted the value in using fast feedback to gain the most accurate responses for free ridership, but it is not required in any state. A number of Illinois utilities use a parallel path evaluation approach for selected custom projects that allows for real-time NTG. In California, pre-screening of custom projects with respect to an initial NTG value to reduce risks of surprise NTG values when the full impact evaluation is performed. This two-step approach helps produce a "no surprise" approach that builds confidence in the NTG estimates.

- **Sensitivity analysis:** Sensitivity analysis (with full transparency regarding the scoring) has been used in all three case study states, primarily when the pre-defined batteries are first developed and tested, but the algorithms are also periodically revisited. This can be important as different but reasonable assumptions used in translating question responses into NTG values can result in different NTG values.
- **Triangulation:** The perspective of vendors is collected in all states for custom projects on a project-by-project basis (e.g., if the customer states the trade ally recommendation was important) and can increase the NTG result. Triangulation with vendors/trade ally surveys is also used to address the influence of factors that program participants may not be well positioned to address—e.g., the relative influence of multiple program influences on program impacts. As noted below, multiple experts noted the difficulty of participants understanding attribution of any individual influence on their decision-making, as there are many potential influences in the EE marketplace.
- **Other best practices*:*** Other best practices mentioned by experts include the following: including multiple factors in the NTG scoring (program influence and other non-program influences), ensuring the questions and weighting are fully vetted, consistency checking, and gaining insight into the project story by spending additional time with the participant to understand the project and possibly meeting with implementation staff knowledgeable about the project.

3. ***Previous program influence and other program influence.*** In Massachusetts, the pre-defined algorithm provides for some credit to be given to previous program influence (i.e., credit for influence that builds over time when a program covers multiple years), while the Illinois and California common batteries do not include this as an improvement to NTG. Experts in those states indicate their general agreement with this policy of assuming that previous program influence results in some current year free ridership. Yet, they mentioned that for this to be fair, spillover studies should fully account for this impact. In terms of other fuel programs (i.e., the influence of gas and electric programs), Massachusetts counts this as a program influence factor, and evaluators in California and Illinois may add other programs as a program influence factor. Experts indicated that it is difficult for a program participant to disentangle the influence of multiple programs (e.g., when more than one entity is providing incentives or information to encourage program participation) and recommended best practices to view simultaneous programs as a single offering for free ridership purposes.[2]
4. ***Market effects and spillover.*** In all three states, experts indicated that there is room for improvement in estimating spillover and market effects. Multiple experts noted that, although assessing free ridership is best done soon after the project, the best time for spillover is at a later point in time. Because of this, there are examples of spillover studies being conducted separately from free ridership studies.

2.3 Massachusetts Case Study

In Massachusetts, the program administrators (PAs) oversee EE programs and the evaluations of those programs. The Massachusetts Energy Efficiency Advisory Council (EEAC) was created by the Green Communities Act of 2008 to guide the development of energy efficiency plans by developing,

[2] The trade allies in the industry may have a better perspective on how overall program impacts may be influenced by multiple programmatic efforts in a jurisdiction.

implementing, evaluating, and monitoring the implementation of these plans.[3] The Massachusetts Department of Energy Resources (DOER) is ultimately responsible for EM&V and provides oversight with support from a team of EEAC evaluation consultants that manage stakeholder processes and the overall EM&V effort. The PAs implement programs consistent with 3-year plans, which are established in collaboration with the EEAC and approved by the Massachusetts Department of Public Utilities (DPU).[4]

Massachusetts has shareholder incentives based on net savings and includes free ridership and participant and nonparticipant spillover in its definition of net savings. The EEAC facilitates a collaborative stakeholder process to define NTG factors on a prospective basis, which are agreed upon for each 3-year program cycle; yet, gross evaluation results based on realization rates are applied on a retrospective basis. Massachusetts uses an agreed-upon, pre-defined algorithm for C&I NTG surveys[5] and uses methods such as sensitivity analysis, triangulation, and a bonus for prior program participation. Table 1 displays basic information about Massachusetts NTG policies.

Table 1. Massachusetts NTG Policies

Structural Components	Massachusetts
Shareholder incentives	Savings and net benefits
Definition of net savings	Includes free ridership, participant spillover, nonparticipant spillover
Application of NTG results	Prospective, including savings claims and target setting
Application of gross evaluation results	Retrospective based on realization rates
Final NTG values	Evaluation studies with review and approval by EEAC with defined appeals process
Frequency of NTG updates	At least every 3 years
Approach to market effects	Allowed to include
NTG bonus for previous program participation?	Yes
NTG uses fast feedback methods	Not required
Sensitivity analysis	Not required for defined algorithms
Triangulation	Included in the standard C&I algorithm
Agreed upon algorithm	Yes[6]

Source: Navigant team research

[3] For additional information on the EEAC purpose, governance, and composition, see the Massachusetts Advisory Council website at http://ma-eeac.org/about/.

[4] For additional information on the EE plans, including the state's EM&V plans, see the Massachusetts Advisory Council website at http://ma-eeac.org/plans-updates/

[5] For the commonly used algorithm and questions, see the Tetra Tech, 2014–*2015 Commercial and Industrial Natural Gas Programs Free-ridership and Spillover Study* at http://ma-eeac.org/wordpress/wp-content/uploads/CI-Natural-Gas-Programs-Free-ridership-and-Spillover-Study.pdf

[6] Tetra Tech, *2014–2015 Commercial and Industrial Natural Gas Programs Free-ridership and Spillover Study*, 2015.

In addition to the document review, the Navigant team interviewed four experts in Massachusetts, including two utility staff, one regulatory staff, and one consultant.

2.3.1 Shareholder Incentives

Massachusetts' shareholder incentives are based on a savings mechanism and a value mechanism, with performance assessed at the portfolio level using cumulative 3-year results.[7] The savings component is allocated to programs on the basis of program dollar of benefits and the value component amount is allocated to programs on the basis of program dollar of net benefits.[8] For 2016-2018, the design-level incentive is set at $100 million for electric efforts and $18 million for gas efforts. Net savings include free ridership and both participant and nonparticipant spillover.[9]

All four experts interviewed reported that shareholder incentives are based on net benefits as a way to align policy goals with PA interests. For example, one expert said, "It makes sense to base incentives on net savings in order to align PA interest with society interest." One expert noted that if the policy goal is to provide value to ratepayers, then net benefits encourage PAs to "perform optimally and minimize cost." Another stated that this incentive mechanism focuses the use of "ratepayer dollars to achieve measurable, attributable savings."

Yet, experts also noted the drawbacks of the shareholder incentive mechanism. First, the use of net benefits was reported by two experts to cause some confusion, as savings achieved by programs must be converted to benefits to find out if the shareholder incentives were earned. Additionally, all experts stated that the use of net benefits puts pressure on evaluation methods to estimate attribution accurately, but there are issues with all NTG estimation methods. One expert noted that, "we are exploring limits of social science to answer questions regarding attribution." This expert continued, saying that there are other fields with social program evaluation where they are concerned with attribution, but "few other fields that have to come up with regularly repeated, highly granular, supposedly precise methods, which is a demanding mission" (see Section 2.3.4).

Several experts noted that if the policy goals were different than maximizing ratepayer value, then using gross savings as a shareholder incentive metric might be appropriate. Three experts mentioned that if the policy goals were total savings (impact on the grid, environmental goals, and carbon or GHGs) then using gross savings as a metric might be more appropriate. Finally, one expert noted that for the purposes of shareholder incentives, using a deemed or negotiated NTG result may be appropriate; however, there still would be a need to conduct research into attribution to inform program design and investment decisions.

2.3.2 Application of NTG Results

Prior to the 2013-2015 program cycle, Massachusetts applied both gross evaluation results (e.g., realization rates) and NTG ratios on a retrospective basis to calculate savings achievements and shareholder incentives. Yet, this caused tension in the system, with significant disagreements over the NTG estimates. It was described by one expert as "really, really bad in the past;" another expert described NTG results as "extremely negative as retrospective tool." A third expert noted, with respect to incentives, that when PAs are "losing money based on subjective studies, it gets ugly." Therefore, in the

---

[7] *2016-2018 Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan*, p 237. http://ma-eeac.org/wordpress/wp-content/uploads/Exhibit-1-Gas-and-Electric-PAs-Plan-2016-2018-with-App-except-App-U.pdf

[8] See gas and electric incentive models at http://ma-eeac.org/wordpress/wp-content/uploads/Appendix-R-Electric-PI-Model.xlsx and http://ma-eeac.org/wordpress/wp-content/uploads/Appendix-R-Gas-PI-Model.xlsx

[9] *2016-2018 Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan*, p 258.

most current cycle, Massachusetts adjusted its policies to apply NTG factors on a prospective basis for each 3-year plan period. The perspective of the MA DPU[10] was summarized in a recent study,[11] stating:

> *The DPU accepted the argument that retrospective application of a NTG ratio creates uncertainty and puts program administrators at risk insofar as they invest in a program with an assumed NTG level that can later be revised downward. The DPU reasoned that this would encourage conservative program planning and implementation that would be unlikely to meet to the aggressive savings goals associated with the Green Communities Act.*

All experts described their views of the tradeoffs related to the prospective versus retrospective use of NTG estimates. Two experts noted that retrospective application of NTG results may be more accurate for estimating actual net savings achievements. Yet, the drawbacks listed by the experts outweighed this positive element. All stated the prospective application of NTG results leads to more effective program planning and more certainty for PAs in terms of the actions needed to achieve program targets and incentives. In addition, all experts reported inherent value in studying net savings and using the results as a planning tool to guide effective spending of ratepayer funds. This information is used to inform program planning and, as stated by one expert, “make rational decisions at the time of investment.” A positive feature noted by experts is that focusing on future planning creates a more collaborative, positive, future-focused environment for stakeholders. They said that it allows the regulators to understand and agree to a prudent use of funds moving forward, and it allows the PAs to use best-available information to develop and implement programs. One expert noted that, “prospective application allowed partnership of EEAC to get best results” (this process is described in Section 2.3.3). Additionally, two experts indicated difficulty with having a framework where gross savings factors are applied retrospectively and NTG is applied prospectively, and expressed interest in a system where all evaluation results are applied prospectively.

Currently, the NTG results[12] are updated prior to each 3-year plan and applied through the duration of the period. Several experts noted this “remove[s] some of the volatility” of NTG results, thereby providing stability and certainty for PAs.

In practice, this means that NTG studies are completed approximately 6 months prior to the start of the plan period and, therefore, must be planned and studied 1-2 years prior to the plan period. For example, for the 2019-2121 plan, NTG studies must be completed by mid-2018 and are being planned and implemented in 2017 and early 2018.[13]

Several experts noted concerns with accepting NTG values for 3-year periods, as it can lead to long lags between study implementation and application of results, especially for the end of the 3-year implementation cycle. One expert stated, “the 3-year lock has been considered great for risk mitigation but not great for perceived accuracy.” Another expert stated that this can lead to risks in fast-moving markets, such as lighting. This expert also noted that, “If [the NTG] estimate is too high, ratepayers take a bath. If it’s too low, the PA stops the program and misses opportunities.” Two experts stated that the 3-

---

[10] Massachusetts Department of Public Utilities, DPU 11-120-A, “Order on Program Net Savings and Environmental Compliance Costs,” August 10, 2012.

[11] Northeast Energy Efficiency Partnerships. Regional Net Savings Research, Phase 2: Definitions and Treatment of Net and Gross Savings in Energy and Environmental Policy. NMR, 2012.

[12] For the detailed list by sector, measure, FR, SOP SONP, and NTG factors are listed, see p. 390 *2016-2018 Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan*. “Appendix B: Net to Gross Impact Factors.”

[13] In this example, the program participants are likely from the 2015-2016 program years.

year lock was more important in the past when the programs were less mature and the NTG results were changing; since then, the results have been more stable.

2.3.3 Process for Finalizing NTG Results

As noted above, in Massachusetts, the DOER is ultimately responsible for EM&V and conducts oversight through a team of EEAC evaluation consultants that manages stakeholder processes and oversight for EM&V. In terms of NTG policies, evaluators recommend NTG results to the EEAC consultants. The EEAC then discusses and debates these results with the goal of gaining consensus on the value, which may be different than the initially recommended result. If there is no consensus, the EEAC consultant makes the final decision, which can be appealed through multiple layers, including the DPU and courts. However, since this system was put in place 8 years ago, there has not yet been a dispute or appeal.

Multiple experts noted the collaborative, transparent process facilitated by the EEAC for finalizing NTG results. Although one expert reported that there are often “multiple meetings and calls to debate results,” the process was described by several experts as collaborative and transparent. One expert noted that because NTG “answers are not easy,” there needs to be a process with “enough room for reasonable people to disagree,” which they felt was provided by the EEAC process.

Several experts noted the value of the transparency and ongoing communication embedded in this process. They referenced the ongoing interaction between EEAC consultants, PAs, and evaluation vendors, which “avoids miscommunication.” This expert stated that, “When utilities have to stay hands-off until the final decision, it leads to issues of misunderstanding the programs and not understanding data appropriately.” Finally, one expert noted that the process was easier than in the past due to the “standardized battery of C&I” NTG questions, which allowed PAs and vendors to “know the rules of the game” and minimize “some of the biggest disputes over results.”

Experts noted that the EEAC often uses evaluation results to determine negotiated values. For example, because the NTG results will be used prospectively, the results may be adjusted to better align with future programs—account changes in future participants, likely changes in the market, or remove outliers from past participants. Experts also agreed that it is important to conduct primary research on NTG, with one stating, “you can’t improve something you don’t understand, so I don’t think you should stop doing the studies. It’s an absolute necessity, at least at a qualitative level, to understand what is changing the world and what is not.”

In terms of access by utilities or PAs to survey data at the individual respondent level, one expert noted that this was not required in Massachusetts and was based on the policies of individual EM&V contractors. In their experience, “contractors will not release information if used to identify specific customers.” However, according to one expert, the evaluation contractors can provide useful information by providing cross-tabulations or frequencies that can be used to understand how responses to certain questions drive the NTG values and conduct sensitivity analyses (e.g., looking at impacts of specific questions on the scoring algorithm). This approach protects confidentiality while providing information necessary to understand what questions and responses affect the final NTG estimates. This has reduced the need to share the detailed data.

2.3.4 NTG Methods

Massachusetts uses multiple methods to estimate net savings, including self-report surveys, market effects studies, econometric (top-down) modeling, quasi-experimental models, and RCTs. When the team asked experts which of these methods are used for custom projects, all answered that only self-report methods are being used. One expert called custom projects the “single best poster child for self-report methods," noting that the custom nature of each project and the markets they work within "rules out other methods," as it cannot easily be randomized or understood econometrically. Therefore, for custom projects, the primary method used in Massachusetts is self-report surveys, and studies use a common but not required battery of C&I questions and NTG algorithm.[14]

Every expert noted the problems with self-report methods for NTG, with one expert calling them "flawed" and another "fallible." One expert referenced Churchill's famous[15] quote "Democracy is the worst form of government, except for all the others" by saying, "Self-report is the worst form of NTG methods, except for all of the others." This same expert noted that self-report methods have drawn a lot of criticism because they are "widely used and affects stakeholders." Yet, they stated that other methods would likely draw "just as many complaints because every NTG method has serious threats to validity." They continued saying that it is "not because of bad choices, but because we are operating at limits of what we can know." Similarly, another expert mentioned that the underlying information—understanding attribution in a counterfactual scenario—is a hard question to answer. Another expert stated that the "only method that is reasonably free from threats to validity is RCT," but also said that it was difficult or impossible to implement for custom programs. It should be noted that the problem of attribution is not unique to EE program evaluation—it is equally challenging for assessing investments and policies across all fields.[16]

2.3.5 Mitigating Issues with Self-Report Methods

The comments of the experts above regarding concerns with self-report surveys make it important to apply methods that can help mitigate the well-documented issues. All respondents mentioned the value of the pre-defined C&I NTG algorithm, which was noted as a "common method but not strictly required or followed," although most studies follow the standard method. Two experts indicated having a common algorithm allows NTG results to be compared over programs and time. Two experts also noted that using a common method allows for the algorithm to be updated, tweaked, and improved over time, leading to a more sophisticated approach that "people are comfortable with because it's been incrementally changed over time."

In terms of specific approaches to mitigating problems with self-report methods, experts had the following insight:

- **Fast feedback:** All respondents noted that minimizing time between when the project is completed and when the NTG survey occurs is important. One stated that staffing can change over time and another "would expect that recall to fade on the issue of freeriding." Yet, one expert reported that minimizing the "lag between intervention and self-report hasn't been a strength" in Massachusetts, and another noted "we would like to have fast feedback, but the MA framework is not fast." Additionally, one expert stated that although fast feedback is best for free ridership, the opposite is true for spillover. Specifically, they said, if surveys "show up too soon, you might not capture all the spillover." See Section 2.4.7 for more on spillover studies.
- **Sensitivity analysis:** Sensitivity analysis was not reported by experts as required or used in Massachusetts; however, it was conducted when the methods were originally developed and is currently being conducted as part of the NTG update process for the next 3-year cycle.

---

[14] Tetra Tech, *2014–2015 Commercial and Industrial Natural Gas Programs Free-ridership and Spillover Study,* 2015.

[15] https://www.winstonchurchill.org/resources/quotes/the-worst-form-of-government/

[16] See Section 2.4 of *Chapter 17: Estimating Net Savings: Common Practices* of *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures.* US Department of Energy, 2014.

- **Triangulation:** At a project level, the algorithm includes vendor perspectives when the respondent reports that the vendor recommendation was an important factor in their decision.[17] This project-specific approach for incorporating trade ally influence works for custom programs due to the size and scope of the projects (i.e., trade allies can recall and comment on specific projects).
- **Other best practices:** Experts also noted other best practices in self-report methods, including time-series check-ins of the NTG within the 3-year plan period to give insight to the implementer of their NTG status and time to make program adjustments in an attempt to improve the NTG. Another noted that the "way questions are stylized is incredibly important."

2.3.6 Previous Program Influence and Other Programs Influence

In Massachusetts, the C&I algorithm has an adjustment[18] for previous program influence, which increases NTG ratios (decreases free ridership). One expert noted the benefit of including this factor as a way to ensure programs gain credit for their long-term efforts with C&I customers.

In terms of multiple programs working concurrently, Massachusetts conducts combined evaluations for its statewide programs and its algorithm specifically accounts for projects that have both gas and electric measures in a comprehensive offering. The algorithm[19] includes questions about the influence of a secondary fuel program; based on this result, it may increase the NTG (decrease free ridership). Several experts noted agreement with this approach, saying that it is not best practice to try to allocate attribution across different program sponsors. Specifically, one expert noting it is "nearly impossible to break out" attribution in a single project and that policies should encourage PAs to work together and provide more comprehensive offerings. Another expert said, there are "so many drivers and influences – utilities, state policies, advertising, stocking behavior, trade allies – so many things going on. We're missing a lot."

2.3.7 Market Effects and Spillover

Although Massachusetts policies allow spillover to be included in the NTG results, several experts stated they think that comprehensive market effects and spillover that are not being captured by the current approaches. Experts reported that there is value in delaying research on spillover to make sure that the influence of the program has had time to impact participant decisions. Another expert noted that they feel that Massachusetts is not fully capturing market effects and they would like to see more efforts in this area.

---

[17] In Massachusetts, the C&I algorithm includes contacting vendors or design professionals identified by program participants as being most influential in their decision to install the electric saving measures through the program and their response can only increase NTG.

[18] Impact of previous participation is calculated through a series of four questions. Participants are asked to state whether they agree or disagree with four statements about the effect past participation has had on their decision-making. Based on the number of statements with which they agree, their free ridership is reduced by 75%, 37.5%, or not reduced at all. Source: *2014–2015 Commercial and Industrial Natural Gas Programs Free-ridership and Spillover Study*.

[19] If a participant rates the influence of the gas project as high (7 or greater on a scale of 0 to 10), the free ridership score remains the same. If the participant rated the influence of the gas project a 6 or less, the free ridership score is reduced by half. According to the Massachusetts free ridership study, this reduction is necessary because the previous factors focus on the specific effect of the program incentive and the overall effect of the program. Without this adjustment, the influence of providing a comprehensive project (one that includes both gas and electric) is understated. Source: *2014–2015 Commercial and Industrial Natural Gas Programs Free-ridership and Spillover Study.*

In terms of market effects, the 2016-2018 plan[20] states that, "to quantify program impacts that have translated to market effects, first a baseline must be established, and then changes from the assumed baseline can be determined to be program induced." Only then can the market effects be counted in net savings.

2.3.8 Other Policies

Massachusetts has an EE resource standard, gas and electric decoupling, and no lost revenue recovery mechanism. Massachusetts has a carbon policy and gross savings were used to set compliance goals with Massachusetts Global Warming Solutions Act (GWSA), as described in the 2016-2018 statewide plan:

> *While [attribution] factors are appropriate for use with the GCA,[21] which seeks to determine which savings resulted from PA program efforts (net savings), the GWSA seeks to quantify all energy efficiency GHG reductions without regard to PA program attribution. Consequently, calculating GHG reductions based upon net savings would undervalue the contribution of energy efficiency to GHG emission reductions.[22]*

2.4 California Case Study

In California, the California Public Utilities Commission[23] (CPUC) is responsible for conducting impact evaluation research on behalf of the utilities, including research into NTG values. The utilities conduct process and market studies in coordination with CPUC staff. The CPUC uses an ex ante review process[24] to review and approve deemed savings estimates and stores measure information, including NTG values, in the Database for Energy Efficient Resources (DEER) and in non-DEER work paper archives.[25]

The CPUC has been responsible for NTG research since 2006, and policies have evolved over that time. For the 2006-2008 period, the financial mechanisms included both shareholder incentives and penalties and were based only on the outcomes of evaluated net savings. During that period, the CPUC conducted the evaluations without input from the utilities but under contract with evaluation consultants that had supported the utilities in the past. In 2009, as shareholder incentives and penalties were being calculated, stakeholders voiced serious concerns about the validity of NTG results and what came to be called the "changing of the goal posts." To improve the transparency and collaboration of the system, the CPUC adopted several changes to the shareholder incentive mechanism and the expectations around cooperation between the utilities and the Commission staff on evaluation activities.[26]

---

[20] *2016-2018 Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan*, p. 254.

[21] Green Communities Act

[22] 2016-2018 Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan P 258.

[23] For additional information on the CPUC's role in evaluation, see the CPUC Energy Efficiency Evaluation Measurement and Verification website: http://cpuc.ca.gov/General.aspx?id=5399

[24] For additional information on the ex-ante process, see the CPUC Ex Ante Review website: www.cpuc.ca.gov/General.aspx?id=4132

[25] Additional information can be found at DEER (Database for Energy Efficiency Resources) http://www.deeresources.com; and Non-DEER work paper web page: http://www.deeresources.com/index.php/non-deer-workpapers

[26] Sangeetha Chandrashekeran; Julia Zuckerman, and Jeff Deason, January 2014, *Raising the Stakes for Energy Efficiency California's Risk Reward Incentive Mechanism*, https://climatepolicyinitiative.org/wp-content/uploads/2014/01/Raising-the-Stakes-for-Energy-Efficiency-Californias-Risk-Reward-Incentive-Mechanism.pdf

Currently, California bases its shareholder incentives[27] on a combination of net savings and three other (non-savings) factors, including custom project review performance, non-resource programs, and codes/standards. The state now applies NTG on a prospective basis for most of the portfolio; however, for inputs and measures that are determined to be "uncertain," such as custom project NTG, the evaluated NTG results are applied retrospectively.[28] The CPUC conducts primary research on NTG values and then uses a regulatory stakeholder input process to update prospective DEER values for NTG. California uses a pre-defined survey for most NTG analysis, including a unique algorithm for residential and C&I that is disclosed in the evaluation planning process.[29] Table 2 displays basic information about California NTG policies.

Table 2. Basic California NTG Policies

Structural Components	California
Shareholder incentives	Net savings (lifecycle) plus three non-savings factors[30]
Definition of net savings	Includes free ridership, participant spillover, nonparticipant spillover
Application of NTG results	Prospective for most of the portfolio for goal setting and claiming savings; retrospective for custom project NTG and other uncertain factors.
Application of gross evaluation results	Same as NTG results process
Final NTG values	Evaluation studies, work paper proposals from utilities, and DEER values
Frequency of NTG updates	Annually, where required by process
Approach to market effects	Separate spillover research
NTG bonus for previous program participation?	No
NTG uses fast feedback	Not required, but pre-screening occurs for custom programs
Sensitivity analysis	Yes
Triangulation	Yes, where possible
Agreed upon algorithm	Yes

Source: Navigant team research

In addition to the thorough document review, the Navigant team interviewed three experts in California, including one utility staff, one former regulatory staff member, and one consultant.

[27] CPUC Energy Savings Performance Incentive (ESPI) http://www.cpuc.ca.gov/General.aspx?id=4137

[28] CPUC Decision 13-09-023 September 5, 2013 *DECISION ADOPTING EFFICIENCY SAVINGS AND PERFORMANCE INCENTIVE MECHANISM* p. 50; http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M076/K775/76775903.PDF

[29] Example of NTG survey instrument available for public review: 2013-14 NTG Standard Very Large Interview Guide https://pda.energydataweb.com/api/view/1199/IALC_Customer_NTG%20survey_Final_11-07-14.pdf

[30] Custom project review performance, non-resource programs, and codes/standards.

2.4.1 Shareholder Incentives

For the 2006-2008 period, California had a shareholder incentive mechanism that was based on net savings and contained both incentives and penalties. Specifically,[31] if utilities met 80% of their net savings goal, they would receive an incentive; if they met less than 65% of their goal, they would receive a penalty. This mechanism was reported by experts to result in disagreements among stakeholders and in large risks in terms of lost incentives based on an uncertain NTG estimate applied retrospectively. One expert said, "When incentives are tied exclusively to net savings, it puts a lot of pressure on evaluation to estimate net savings with a high degree of accuracy, which the science cannot provide."

When the evaluations for the 2006-2008 period were completed, they revealed that the goals were not achieved. One expert said that, "the entire process was brought into the political realm and policymakers were, among other things, frustrated that the utilities and the CPUC staff and consultants were not able to resolve their own disputes through collaborative engagement." In a 2010 decision, the Commission made several modifications to the original rules for the incentive mechanism and awarded the utilities financial incentives based on the changes.[32] The Office of Ratepayer Advocates filed a petition in 2014 (note this was 6 years after the end of the program cycle being evaluated) requesting the CPUC rescind the payments based upon improper communications between a utility and a Commissioner in violation of state law and CPUC rules. In 2015, the CPUC issued a decision re-opening the case, which was ultimately resolved via settlement approved by the Commission in October 2016.[33]

Beginning in 2013, after thorough reconsideration of a revised incentive mechanism, the CPUC adopted the Energy Savings Performance Incentive (ESPI) mechanism, which provided roughly $25 million in ex ante and ex post incentives in 2015 for net savings accomplishments. The incentive payments are based on achievements against four separate factors:

1. Net savings, which is calculated separately for certain and uncertain measures (see Section 2.4.2 for more detail). Net savings includes free ridership and participant and nonparticipant spillover, although California uses an estimated 5% adder[34] for spillover to portfolio savings.[35]
2. Ex ante review performance, which represents effectiveness of utilities in implementing the pre-review of custom projects, such as being timely and having proper documentation.
3. Codes and standards, based on utility expenditures for codes and standards advocacy, compliance, and other program activity.
4. Non-resource programs, based on utility expenditures for programs that do not achieve direct energy savings.

---

[31] Described in detail in Sangeetha Chandrashekeran; Julia Zuckerman, and Jeff Deason, *Raising the Stakes for Energy Efficiency California's Risk Reward Incentive Mechanism*, January 2014, https://climatepolicyinitiative.org/wp-content/uploads/2014/01/Raising-the-Stakes-for-Energy-Efficiency-Californias-Risk-Reward-Incentive-Mechanism.pdf

[32] Decision 10-12-049; DECISION REGARDING THE RISK/REWARD INCENTIVE MECHANISM EARNINGS TRUE-UP FOR 2006-2008; December 16, 2010. See ordering paragraphs. http://docs.cpuc.ca.gov/word_pdf/FINAL_DECISION/128879.pdf

[33] Office of Ratepayer Advocates Review of 2006 - 2008Risk / Reward Mechanism http://ora.cpuc.ca.gov/general.aspx?id=3626

[34] CPUC Decision 13-09-023 September 5, 2013; *DECISION ADOPTING EFFICIENCY SAVINGS AND PERFORMANCE INCENTIVE MECHANISM,* p. 27.

http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M076/K775/76775903.PDF

[35] CPUC Decision 13-09-023 September 5, 2013; *DECISION ADOPTING EFFICIENCY SAVINGS AND PERFORMANCE INCENTIVE MECHANISM* Ordering Paragraph 3, p. 94.

http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M076/K775/76775903.PD

Experts generally reported that the current mechanism has been less controversial than the previous mechanism, with one saying, "it is better now because the risk of NTG results have been mitigated [with the use of multiple factors in the incentives calculation] and the process is more transparent." The utilities are also more consistently using the NTG ratios the Commission recommends in their program planning and implementation; therefore, less divergence is occurring in savings claims and evaluated results on NTG. In terms of multiple factors, one expert noted that the new incentive mechanism "recognizes that the portfolio is diverse and recognizes that the Commission is asking utilities to do more than just acquire savings." The current incentive mechanism was described by an expert as a "reasonable compromise"—the mechanism is clear about which and how much of the incentive payments are at risk based on measure or program performance (see Section 2.4.3). One expert said that, "Shareholder incentive is nice – but people aren't thinking about it much anymore," which they attributed to the relatively small size of the incentive payment compared to the previous mechanism.

The incentive mechanism in California is tied to the lifecycle net savings, given that goals are currently net savings and the incentive structure is intended to encourage long-term investments in EE. Although goals were gross savings during the 2009-2012 period,[36] net goals have more recently prevailed as better aligned with how EE is included in the state's load forecasting.[37] Additionally, a recent potential study estimated net potential above naturally occurring adoption (free ridership).[38]

2.4.2 Application of NTG Evaluation Results

During the 2006-2008 period, California applied all NTG and other evaluation results on a retrospective basis for the incentive payments. As noted above, this created a system that caused friction among stakeholders because of the perception of uncertainty and the implications of the results. In particular, stakeholders voiced concerns about "moving the goal posts," meaning that the expectations around judging performance appeared to change (related to retroactive application). One expert said, "retroactive application of results were very problematic, especially when a decision is made after programs have been designed and implemented, it messes with the business planning structure."

Therefore, in addition to adjusting the stakeholder incentive as described above, California also adjusted its framework to move NTG and other evaluation results to a prospective basis. The current California model has an annual process to update as many deemed savings parameters as possible with new information coming from recent evaluations studies and other factors like code updates or major market trends. On this schedule, the majority of the portfolio inputs (such as savings and NTG values) are updated on a prospective annual basis and used for program planning and goal setting.[39] Each year's measures and parameters in the portfolio for the upcoming year are reviewed for their contributions to "uncertainty." If they are selected, then payment for energy achieved for those measures will be delayed (or applied retrospectively) for the purposes of the incentive payment. This structure, according to one expert, is designed to hold out some portion of incentive payments until the uncertainties are resolved

[36] For more information, see CPUC, "History of California Public Utility Commission Goals for Energy Efficiency," 2010. www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=4069

[37] California Energy Commission, *ESTIMATES OF ADDITIONAL ACHIEVABLE ENERGY SAVINGS: Supplement to California Energy Demand 2014 - 2024 Revised Forecast,* SEPTEMBER 2013 CEC- 200- 2013- 005- SD; p. 9. http://www.energy.ca.gov/2013publications/CEC-200-2013-005/CEC-200-2013-005-SD.pdf

[38] Navigant Consulting, Inc. *Energy Efficiency Potential and Goals Study for 2018 and Beyond, Final Public Report,* Prepared for: California Public Utilities Commission, September 25, 2017. (http://www.cpuc.ca.gov/General.aspx?id=6442452619)

[39] CPUC D.15-10-028 October 22, 2015, *DECISION RE ENERGY EFFICIENCY GOALS FOR 2016 AND BEYOND AND ENERGY EFFICIENCY ROLLING PORTFOLIO MECHANICS*; Section 3.2.4. Rolling Portfolio Cycle Schedule http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M155/K511/155511942.pdf

through ex post evaluation. The list of uncertain measures is developed each year and reviewed by stakeholders (see Section 2.4.3). C&I custom programs are always included in this list based on the Commission rules due to their variability based on the custom nature of participation.

Experts noted that, although a portion of the portfolio remains retrospectively applied, the system now works much better because these items are “called out in advance and utilities have a chance to weigh in” at the time of that decision and provide feedback. Additionally, one expert stated that NTG results have “changed little from year to year for the custom programs, so all parties and stakeholders understand how they will be evaluated, and can manage for successful outcomes in their evaluations.”

For custom projects, the retrospective application of results has led to some utilities employing pre-review and screening of individual projects to assess NTG and baselines prior to project approval. This helps the utilities decide if they want to invest in large individual projects, with prior knowledge of the likely result of the NTG study. One expert indicated this has led to changes in program design (e.g., eligibility requirements) to try to minimize free ridership.

2.4.3 Process for Finalizing NTG Results

For the 2006-2008 program cycle, CPUC staff conducted evaluations without utility involvement (except to provide datasets) because of the strict legal construct for the shareholder incentive payments in place at the time, according to one expert. Results were presented as final for the purposes of determining the eligible incentive amounts. Given the conflict that resulted, the Commission required adoption of a more cooperative structure for evaluation processes in the next program cycle (2010-2012).

In addition to submitting a joint master evaluation plan,[40] the primary change in the processes was that all evaluations would be conducted with key points for public engagement. Commission staff developed a process that reflected this requirement in three basic steps, which is still currently used (see Table 3):

[40] *2013-2016 Energy Division & Program Administrator Energy Efficiency Evaluation, Measurement and Verification Plan Version 7*, available at: https://pda.energydataweb.com/api/view/1688/EM&V%20Evaluation%20Plan%202013-2016%20Plan%20V.7%20December%202016-lastround-12-5-2016.pdf.

Table 3. California Evaluation Review Process

General Expectation	Process to Meet Expectation
Specify what will be evaluated (for purposes of the incentive payment and generally)	• Publish an annual master evaluation plan (CPUC and investor-owned utility [IOU] staff) • Identify evaluation priorities for specific sectors and estimated budgets • Solicit public input on high level priorities
Publicly vet evaluation plans	• Evaluation plans for CPUC and IOUs posted for public comment • Most study plans also have a webinar to discuss evaluation priorities and methods
Publicly vet results for comment prior to finalization	• Studies are shared in draft form for public comment • Deviations from the methods in the evaluation plan are highlighted • Implications of results (for incentive payment) or for DEER updates are highlighted
(extra step) Response to recommendations	• Program implementers are required to respond in writing as to what actions they will take on the recommendations in the report

Source: Section 5.4 Energy Division IOU Collaboration in the Master Evaluation Plan

All experts stated this level of transparency was an improvement and has led to substantial reduction in conflict over evaluation results.

Prospective savings estimates and other parameters are developed in the DEER process.[41] Annually, certain measures or baseline assumptions in the DEER database are updated using new evaluation information or other market information (like new codes and standards). Based on this information, the DEER team recommends savings and NTG results to be used in the upcoming program years. These values may differ from the evaluation results because of the need to forecast, but the differences from evaluation results are explained and justified by the DEER team. The proposed updates are shared for public review and stakeholders provide input, which may drive further changes. The Commissioners approve the values either in a ruling or a decision, but largely CPUC staff, consultant, and stakeholders attempt to resolve disputes in advance.

The uncertain measures process (for retrospective applications) is conducted annually. In the fall of each year, the CPUC hosts public meetings for any stakeholder to weigh in on the evaluation priorities for the year in the master evaluation plan. In addition, an uncertain measure list is published each year identifying the technologies and parameters that must be evaluated for stakeholder incentives to be made on the savings claims.[42] As specific evaluations are planned, the evaluation plan is posted on the CPUC's public comment platform and a webinar is typically held to solicit feedback on the scope of the evaluation and even specifics of the methods. This includes how NTG will be determined and sharing survey instruments and algorithms that will be used to assess free ridership or spillover. At the end of the evaluation, the draft results are shared publicly. This allows all stakeholders to weigh in on the results and confirm that the agreed-upon methods were followed.

[41] CPUC Ex Ante Review website: www.cpuc.ca.gov/General.aspx?id=4132

[42] 2018 Final Uncertain Measure List; http://www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=6442455469 or http://www.cpuc.ca.gov/General.aspx?id=4137

One expert noted that if utility representatives want to see individual results for large custom projects, then CPUC staff and consultants will hold a meeting with utility representatives because these may contain confidential participant information. After comments are processed, the evaluation consultants finalize the report and post it publicly.[43] The utilities are required to respond to the recommendations in the report within 60 days and the results are used to inform their incentive payment claims. If they make an incentive claim that is counter to the evaluation findings, they have the burden of proof to demonstrate those results are wrong; however, that has not happened yet. The Commission makes the final determination on the incentive payment by the end of each year.

2.4.4 NTG Methods

In California, multiple methods are used to estimate net savings, such as self-report, econometric choice models, and market surveys. One expert noted that "market approaches" to understand NTG are better than self-reports, but there are challenges to this because there must be both treated and untreated markets and the result is a combination of free ridership and spillover together. Another expert stated that there are only a "limited pool of NTG tools that can be deployed after the program has happened" and that "better tools require more integration with program design and can be more expensive." Two experts discussed the potential opportunity to use industry standard practice baselines as a potential alternative to NTG, but also noted the difficulties with this method, such as understanding industry practices for custom projects.

Therefore, the experts reported that self-report surveys using pre-defined questions and algorithms are used as the primary method for estimating free ridership. One expert stated that "methods must be sensitive to how the program is deployed" and, therefore, self-report surveys make sense for custom projects. This is because "deep conversation with [a] customer seems reasonable," but that it needs to be backed up with documentation of how the utility or program implementer intervened in the decision-making process to lead to the more efficient outcome.

For C&I, a consistent set of questions is used for determining NTG.[44] It was developed to improve consistency and transparency by using a consistent survey instrument. One expert mentioned the value in having an approach that is "reliable year to year," thereby providing stability in approaches. This has led to lower variance in results and more stability in NTG values. However, this expert also noted, "consistent NTG values doesn't mean it's accurate." This same expert stated they have "zero faith in any self-report method," as it is asking participants counterfactual questions that they "probably don't know the answer to but they will give an answer anyway."

2.4.5 Mitigating Issues with Self-Report Methods

Having conducted evaluations for years that leverage self-report methods, California has developed and implemented multiple approaches to mitigate issues with these methods, including:

- **Fast feedback:** California does not have specific policies on the timing of NTG surveys. As noted above, because of the retrospective application of results for shareholder incentives, some utilities (and the Commission) are using pre-screening on NTG prior to project approval.

[43] Example of Final Impact Evaluation Report for Commercial Industrial and Ag: https://pda.energydataweb.com/api/view/1845/IALC%202015%20Custom%20Report%20Final.pdf

[44] The Nonresidential Net-To-Gross Ratio Working Group, *Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers*, Prepared for the Energy Division, California Public Utilities Commission, October 16, 2012.

- **Sensitivity analysis:** Sensitivity analysis is required in the California algorithm.[45] One expert stated that although they conduct "sensitivity analysis and scenarios for every evaluation, it makes little difference" because the algorithm has been tested and refined over time.
- **Triangulation:** Experts mentioned that program-level triangulation is occurring for residential programs, especially for upstream programs where there are multiple streams of information. For custom projects, California uses vendor surveys to assess program influences on their customer recommendations. However, this score is only used if the participant rated the vendor score as the highest influence and the vendor indicated the program was highly influential. In these cases, the vendor score increases the NTG; in no instance would it decrease the result.
- **Multiple scores:** In California, the NTG questions and algorithm includes both program and non-program (e.g., corporate policy) influence scores to account for different ways of measuring program influence.
- **Other best practices:** One expert had multiple suggestions for mitigating issues with self-report surveys, including using warm-up questions to improve the discussion, understanding the project story, delaying the counterfactual questions until later in the discussion, and ensuring that consistency checks happen before an interviewer gets off the phone. The same expert indicated that they do not think that short surveys of NTG are accurate for custom projects, as they do not fully account for the unique nature of these projects.

2.4.6 Previous Program Influence and Other Program Influence

The California NTG framework does not give additional NTG credit for previous program influence. One expert stated that the algorithm is focused on measuring influence within the period and it does not "look backward." Although this expert described it as "possibly a little harsh," they also said that there would have to be a limit on counting previous program influence, saying, "How do you decide how far you go back on program influence?" On the other hand, one expert noted that California does account for cases in which a technical assessment or audit was done a few years ago, which is considered as current program influence.

In California, joint program participation could be accounted for in the program components score. The question lists a variety of possible program and non-program influences and asks the participant to rate the influence of that aspect on their decision. In a program where the joint fuel nature was important, the evaluator can add this to the list of possible program influence.

2.4.7 Market Effects and Spillover

Experts noted the importance of capturing spillover to understand program influence, but also noted the difficulty in estimating spillover. California has a unique approach to NTG in that it includes free ridership on a project or program basis but includes spillover using a portfolio-level adjustment. This adjustment is meant to account for both participant and nonparticipant spillover, and was set at 5%, based on the fact that there was no credible estimate for the actual amount of spillover and applying a percent adjustment to the whole portfolio was easier than estimating per measure.[46] The CPUC stated,[47] "We find it more

---

[45] *Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers,* October 2012.

[46] Decision 12-11-015; November 8, 2012 *DECISION APPROVING 2013-2014 ENERGY EFFICIENCY PROGRAMS AND BUDGETS*, pg. 55 http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M034/K299/34299795.PDF

[47] *DECISION APPROVING 2013-2014 ENERGY EFFICIENCY PROGRAMS AND BUDGETS*, p. 55.

appropriate to apply a portfolio-level 'market effects adjustment' of 5% across the board to the entire 2013-2014 portfolio cost-effectiveness calculation in recognition that California's long history of commitment to energy efficiency resources has resulted in measure adoption outside of program channels."

Recently, the CPUC completed a spillover study that found some differences among sectors (residential was higher than C&I) and by fuel type.[48] One expert indicated that their perspective on this study was that "spillover is even harder to quantify and estimate" than free ridership. Another expert noted that a "separate spillover interview is probably best practice" and that delaying the interview will "get the best spillover insight."

2.4.8 Other Policies

California has specific climate change and carbon reduction policies: AB 32 (California Global Warming Solutions Act of 2006[49]) and more recently, SB350 (Clean Energy & Pollution Reduction Act of 2015[50]). Both raise expectations for meeting and increasing EE goals. When asked about the relationship between carbon policies and net savings, one expert said the discussion is similar to the gross versus net conversation. On the one hand, overall carbon reduction (and consumption reductions) is the focus of carbon policies. Yet, the additional effect over what is already happening in the market (i.e., net savings) is also important to understand the incremental impact of activities. The same expert opined that EE has never been fully reconciled with carbon goals. Assumptions for deemed savings measures are not tightly aligned with realistic avoided emissions. With a greener grid in California, the value of being leaner through efficiency will be more dependent on time and location than it has in the past.

2.5 Illinois

In Illinois, PAs[51] are typically gas and electric utilities, and they are responsible for managing evaluations[52] with oversight from the Illinois EE Stakeholder Advisory Group (SAG). SAG[53] reviews EE plans including portfolio and program designs, draft EM&V workplans, and the Illinois Technical Resource Manual (TRM).[54]

---

[48] Opinion Dynamics, *PY2013-2014 CALIFORNIA STATEWIDE RESIDENTIAL AND NONRESIDENTIAL SPILLOVER STUDY*, 2017. https://pda.energydataweb.com/api/view/1936/CA%20Statewide%202013-14%20Res%20Nonres%20Spillover%20Report%20DRAFT%202017-08-18%20(2).pdf

[49] See 2006 Assembly Bill No. 32, Chapter 488 at: http://www.leginfo.ca.gov/pub/05-06/bill/asm/ab_0001-0050/ab_32_bill_20060927_chaptered.pdf

[50] See 2015 Senate Bill No. 350, Chapter 547 at http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB350

[51] Typically, PAs are utilities: Ameren IL, ComEd, Nicor Gas, and Peoples Gas-North Shore Gas.

[52] For more information on roles of various parties, please see: Illinois Energy Efficiency Policy Manual Version 1.1 - A Manual Guiding the Operation of Illinois Energy Efficiency Programs. See: http://ilsagfiles.org/SAG_files/Subcommittees/IL_EE_Policy_Manual_Subcommittee/2017_Revision/IL_EE_Policy_Manual_Version_1.1_5-5-17_FINAL.pdf.

[53] For more information on the SAG, see IL SAG website: www.ilsag.info/net-to-gross-framework.html and Illinois Energy Efficiency Policy Manual Version 1.1 - A Manual Guiding the Operation of Illinois Energy Efficiency Programs. http://ilsagfiles.org/SAG_files/Subcommittees/IL_EE_Policy_Manual_Subcommittee/2017_Revision/IL_EE_Policy_Manual_Version_1.1_5-5-17_FINAL.pdf. Note the NTG framework has been superseded by the NTG policy in the Policy Manual.

[54] See current and historic IL TRMs: http://www.ilsag.info/technical-reference-manual.html

Illinois does not currently have shareholder incentives or penalties, although the electric utilities will have shareholder incentives beginning in 2018 and gas utility annual energy savings goals are adjusted to align with changes to Illinois TRM values.[55] Illinois began using a prospective approach to NTG factors in 2013, including all areas of the portfolio. SAG oversees the collaborative stakeholder process that defines the prospective NTG factors, which includes free ridership and participant and nonparticipant spillover. Illinois' pre-defined NTG questions and algorithms are included in the TRM, which includes unique variations for residential and C&I programs.[56] Table 4 displays basic information about Illinois NTG policies.

Table 4. Basic Illinois NTG Policies

Structural Components	Illinois
Shareholder incentives	None until 2018
Definition of net savings	Includes free ridership, participant spillover, nonparticipant spillover
Application of NTG results	Prospective, including savings claims and targets
Application of gross evaluation results	Custom is retrospective using realization rates; TRM-based measures are prospective
Finial NTG values	Evaluation studies, SAG process
Frequency of NTG updates	Annually
NTG bonus for previous program participation?	No
NTG uses fast feedback	Not required, increasingly used for selected programs
Sensitivity analysis	Yes, required by TRM
Triangulation	Yes
Agreed upon algorithm	Yes, in TRM

Source: Navigant team research

In addition to a thorough document review, the Navigant team interviewed three experts in Illinois, including two evaluators and one regulatory staff.[57]

2.5.1 Shareholder Incentives

As noted above, Illinois does not currently have shareholder incentives; however, in 2018, the electric utilities will have shareholder incentives based on net savings. One expert reported that "in an ideal world" the shareholder incentives would be based on multiple factors with "the bulk of money on benefits and some amount of money on the other policy objectives you care about" such as comprehensiveness and geographic equity. Gas utilities have historically had penalties (and no shareholder incentives); while

[55] Illinois EE Policy Manual Version 1.1. Section 6.2, Adjustable Savings Goals

[56] The most recent version of the NTG questions and algorithms can be found in Volume 4 of the IL TRM v6.0 (http://ilsagfiles.org/SAG_files/Technical_Reference_Manual/Version_6/Final/IL-TRM_Effective_010118_v6.0_Vol_4_X-Cutting_Measures_and_Attach_020817_Final.pdf)

[57] An interview with a utility representative was scheduled, but interviewee was unable to attend due to personal circumstances.

this remains the same, beginning in 2018, their annual energy savings goals will be adjusted to align with changes to Illinois TRM values.[58]

Although Illinois has historically not had shareholder incentives, the use of net savings has been influential in the state. It is used to calculate savings achievements, which could have resulted in penalties for savings shortfalls. Experts noted that using net savings is valuable to ensuring that the impact of utility programs are monitored because, as one expert noted, “utilities have a decent amount of influence in terms of how they influence programs to push higher NTG or lower.” Another expert stated that they “think it’s a good idea to get utilities invested in doing a good job and helps to build support senior executives.” A third expert mentioned that assessing net savings is particularly important for custom programs because it is common to pay for projects that would have happened otherwise. The best programs influence the market by “marketing, technical assistance and identify opportunities and convince customers to do it.”

The Illinois TRM is consistent with the experts the team interviewed in terms of the importance of attribution, while also expressing caution with the ability to know the counterfactual, stating:[59]

> *Attribution provides credible evidence that there is a causal link between the program activities and the outcomes achieved by the program. Attribution research estimates the difference between the outcomes and those that would have occurred absent the program (i.e., the counterfactual). As such, it is important to realize that the concept of the counterfactual cannot be proven with certainty. This statement is not about poor methods, but about the counterfactual itself. Because programs work with people and are usually not a laboratory experiment that can be replicated over and over to find out what actions people would have taken absent an intervention, one would need a time machine to take people back in time and not provide the program. Since time machines do not exist, evaluators have developed methods that approximate the counterfactual to the best of their ability.*

2.5.2 Timing of Evaluation Results

Historically, Illinois had a system with retrospective application of gross and net savings factors. Experts generally agreed that the uncertainty this caused utilities was not productive. One expert noted that the value of retrospective NTG was that it measured “actual performance,” but they said structure caused utilities to “worry about it and become risk averse,” which could lead to them doing a sub-optimal job in delivering programs. One expert said that utilities wanted to be able to track whether they were meeting their goals, but they could not manage the risk “if they were unsure what the NTG value would be.” In fact, another expert noted that this system had “risk to the utility with no means for them to react to the uncertainty.” They continued, saying the policy, “didn’t align incentives with producing better savings or better programs” and discouraged innovation, as utilities were trying to minimize NTG risk.

In 2013, Illinois moved to a prospective approach to its NTG results and since 2016, NTG algorithms have been included in the TRM. The prospective NTG values are produced annually by independent evaluators, reviewed by stakeholders, and finalized by October 1 of each year. These new NTG values are prospectively effective 3 months later to begin the calendar year.[60]

---

[58] Illinois EE Policy Manual Version 1.1. Section 6.2, Adjustable Savings Goals

[59] IL TRM v.60, Volume 4, p 22 (http://ilsagfiles.org/SAG_files/Technical_Reference_Manual/Version_6/Final/IL-TRM_Effective_010118_v6.0_Vol_4_X-Cutting_Measures_and_Attach_020817_Final.pdf).

[60] Illinois Energy Efficiency Policy Manual Version 1.1

Experts noted that early in the process to move to prospective values, there was concern that utilities would have less incentive to monitor NTG and improve their programs; however, because the results are reviewed annually, they are seeing that programs have an incentive to minimize NTG. Three experts stated it is important in a prospective framework to update NTG results regularly. One expert also mentioned that evaluations conducted annually on each program may expend too many resources, so finding a balance of keeping NTG up-to-date but reducing costs is important.

The experts' views on the value of the prospective approach were consistent with documentation in the Illinois NTG framework.[61] It discusses several benefits of the prospective framework for NTG:

- Higher certainty of claims for PAs in terms of meeting targets and achieving incentives—reduces short term performance risk.
- Provides a strong, albeit diminished, incentive for PAs to work to maximize NTG ratios and net savings by continually doing the necessary research to understand markets and make program changes as appropriate in a timely fashion.
- Ensures that decisions about new initiatives or significant program changes are made, recognizing and balancing performance risk as part of the overall portfolio. This provides PAs with an incentive to design and deliver these programs to minimize free riders initially. Thus, PAs can experiment with innovative strategies.

2.5.3 Process for Finalizing on NTG Results

In Illinois, SAG conducts an annual process to update the prospective NTG values and then updates them in the TRM. Annually, evaluators recommend prospective NTG values based on evaluation results and include the reasoning for their recommendation. This is presented to the SAG, where members have the opportunity to "question, challenge and suggest modifications to the evaluators' initial recommended deemed NTG Ratios for the upcoming program year." [62] The goal of the meeting is to reach consensus on the values. If consensus is not reached, then the SAG facilitator develops a document that identifies the issue, different opinions, and the basis for those opinions and then shares with SAG participants an opportunity to review and comment. If consensus still cannot be reached, the evaluators' final recommended NTG values are used. These values may be different from the evaluators' initial recommendation because they consider all comments and discussions from the SAG process.

Experts reported that the current process is smooth and collaborative. One expert said that early in the process there were "intense weeks and months" of debating and negotiating on NTG. Another expert said, "there was a fair amount of SAG values that were not evaluation recommendations," as varying results (e.g., gas and electric) led to negotiated results. However, over time, the process has become faster and evaluation results are being used more often. One expert said the "process is down to a few meetings to determine future NTG."

In Illinois, they also have a process for a mid-year determination of NTG values for new programs that come up during the year. The process is the same: evaluators recommend a NTG result, but instead of actual results, this "may be based on secondary research, when that research produces relevant results,

[61] Optimal Energy, *Proposed Framework for Counting Net Savings in Illinois*. 2010. Available at: http://ilsagfiles.org/SAG_files/NTG/NTG%20Framework.pdf.

[62] Illinois Energy Efficiency Policy Manual Version 1.1.

…otherwise a NTG Ratio of 0.80 will be used."[63] One expert described this as "helpful and has provided flexibility for mid-year program components that come up."

In terms of data access to individual results, evaluators manage the confidentiality of the data. Though, one expert stated that it is possible that evaluators would provide the Illinois Commerce Commission with individual results if requested, and NTG data has been shared with the utilities to conduct their own sensitivity analysis.

2.5.4 NTG Methods

The Illinois TRM lists several methods for NTG estimation,[64] including self-report, econometric/revealed preference approach, RCTs and quasi-experimental designs, deemed or stipulated NTG ratios, market analyses, structured expert judgment approaches, program theory-driven approach, and case studies design. The TRM notes that several of these are not used in Illinois, including common practice baseline approaches and market analyses.

As discussed in Section 2.3.1, the TRM notes concerns with self-report methods of the counterfactual. One expert states that they think self-reporting is a "terrible method but better than all the rest" and that they "would love the ability to do more RCTs" to estimate net savings. This expert noted skepticism of Delphi approaches and common practice baselines, which they see as not appropriate to measure program influence.

Although the TRM lists many options, Illinois evaluations primarily uses self-report methods for NTG estimation. For custom projects, experts did not report any other approaches than self-report to estimate net savings for custom projects. One expert said, "by definition, custom is different kinds of products, so you really have to talk to customer."

Prior to 2015, there were multiple variations of self-report surveys and algorithms being used in the state, which were noted in 2014 to be causing uncertainty as, "different evaluation methodologies, contractors, and simple random statistical variation can influence the measurement of NTG, resulting in a higher than desired level of uncertainty for PAs if used solely on a retroactive basis." [65] This document recommended that:

> *Wherever possible, joint and consistent statewide evaluations be performed. This will eliminate these uncertainties, allow for more direct comparison between PA's performance, as well as provide economies of scale and greater consistency and certainty to PAs about likely future evaluation results. We propose that standardized approaches to measuring free ridership and spillover be adopted in Illinois that ensure consistent measurement both across territories and over time.*

This led to Illinois Commerce Commission orders[66] for each utility that required the developing and adopting consistent statewide NTG methodologies (IL-NTG Methods). For example, the Nicor docket noted it would "help ensure the independence of the evaluators, to improve efficiency in the evaluation process, and to ensure programs across the state as delivered by the various Program Administrators

[63] IL TRM v.6, Volume 4

[64] IL TRM v.6, Volume 4, p. 91-95.

[65] Optimal Energy, *Proposed Framework for Counting Net Savings in Illinois.* 2010.

[66] For example, see Nicor Gas Order (Docket No. 13-0549) – p 41/42.

can be meaningfully and consistently evaluated." It also noted that the "adoption of IL-NTG Methods would save on ... limited evaluation resources by having a common reference document for the evaluators to use in estimating net savings for Illinois."

Based on these orders, SAG undertook a process to develop NTG methods for inclusion in the statewide TRM. These were first included in Illinois TRM v5.0 and contain differences for sectors and programs.[67]

All experts noted the benefit of the common methods, with one saying, "contentiousness was reduced because differences in methodologies were mitigated." Another said it "mitigates debate about which method is used and there's a designated time for that debate to occur" among stakeholders. This expert continued, saying that the common method "mitigates uncertainty and mitigates potential for gaming." Another expert said that the process of determining prospective NTG values is much easier with "everyone using the same approach, because there is less room for differences in approaches," and arguing is reduced because "if numbers are different it's hard to argue due to methodology."

The TRM is updated annually,[68] and SAG reviews any changes to it. The TRM describes the NTG methods as "partially binding," but it allows evaluators to deviate from the methods if it submits a proposal to SAG and gains their approval.[69]

2.5.5 Mitigating Issues with Self-Report Methods

Illinois mitigates issues with self-report methods through multiple approaches:

- **Fast feedback:** Illinois does not have specific policies on the timing of NTG surveys. Several experts noted that fast feedback is better for free ridership accuracy, with one indicating that programs should "measure free ridership as soon as possible after the decision." Another expert said fast feedback is "best practice – if a lot of time has passed, the customer will internalize the decision, or they forget." Due to the annual updates of NTG values in Illinois, several experts mentioned that increasingly NTG research is being conducted in waves throughout the year, on a rolling basis soon after projects are completed.
- **Sensitivity analysis:** Sensitivity analysis is required in the Illinois algorithm for C&I measures. This was particularly important as the algorithms were first being developed, and it was difficult to reach consensus on the scoring approach; thus, the TRM actually required multiple methods for scoring to be presented as part of the prospective approval process. After initial analysis, however, sensitivity analysis has become less important. One expert called it a "good idea" but said "no one pays attention to it," as the results and algorithm have been tested over several years and have stabilized.
- **Triangulation:** Experts reported that program-level triangulation is used for some programs, particularly for residential. The Illinois TRM does also propose using a weighted scoring method to integrate results from various perspectives based on perceived accuracy, data reliability, and statistical confidence/precision. The weight for each method is the average score for that method divided by the sum of the average scores for all methods. For custom projects, however, the

[67] Various NTG for: core non-residential (free ridership and spillover) protocols, free ridership only for: small business, C&I new construction, study-based and technical assistance. For residential, there are cross-cutting approaches and program-specific ones such as upstream lighting, prescriptive rebates, new construction, multifamily, and home energy audits. Across sectors includes behavioral protocols (including RCTs and non-randomized designs). See IL TRV v5.0, Volume 4.0.

[68] IL TRM v.6, Volume 4, Updating the IL-NTG Methods

[69] IL TRM v.6, Volume 4, Procedure for Non-Consensus Items

Illinois TRM vendor influence is accounted for in the program components score—although it could be considered either a program or non-program influence. If a vendor is considered a program influence, a high rating from the participant would trigger a vendor interview and program influence on the vendor would be assessed through a separate survey, which could increase the NTG result.[70]

- **Multiple scores:** In Illinois, the NTG questions and algorithm includes multiple factors to account for different ways of measuring program influence. For C&I programs, the Illinois algorithm averages the program factor, program influence, and no-program scores,[71] while for residential, only the program influence and no-program scores are used.
- **Other best practices:** Experts also discussed other best practices to mitigate issues with self-report surveys, including: "not relying on a single response," probes for contradictions, and "indirect questions that give a sense of whether customers would have done this or not." Another expert stated that it is important for custom projects to use professional interviewers in order to understand the project story, "follow-ups to avoid non-response bias," and making sure "question wording is good."

2.5.6 Previous Program Influence and Other Programs' Influence

In the Illinois TRM, previous program participation is not directly accounted for and does not give a bonus to NTG. Joint program participation in Illinois may be accounted for in the program components score, as the question lists a variety of possible program and non-program influences and asks the participant to rate the influence of that aspect on their decision. In a program where another joint fuel program was important, the evaluator can add this to the list of possible program influences. One expert noted that asking customers to understand unique influences of various programs is difficult, saying, "equipment is a single decision, not multiple decisions…To ask customers to tease influence is not realistic." Another noted that the best practice in joint programs is to conduct a "single evaluation and single NTG value."

2.5.7 Market Effects and Spillover

The Illinois Policy Manual is clear that free ridership should be included in all NTG ratios, but it is somewhat vague on spillover, saying:

> *Spillover shall be included whenever possible and feasible in each NTG calculation. Whenever a NTG value is calculated for components of a Program, it will still include Free Ridership, and if feasible, Spillover. … Evaluators are not required to always include Spillover in NTG calculations due to the costs of Spillover research, but excluding Spillover might unfairly reduce Program calculated savings. Evaluators should consider Spillover, including logical reliance on deemed values and secondary research developed from evaluations of other Illinois Programs and other jurisdictions, to estimate Spillover in relation to the predicted impacts of such Measurements. Also, a sector or Portfolio-level Spillover analysis should be considered by each utility at least once every Plan period when it is feasible and considered viable by evaluation. All such Spillover research should be conducted while being mindful of costs and other evaluation needs.*[72]

---

[70] IL TRM v.6, Volume 4, p. 28-33.

[71] The program factor is the maximum importance of the incentive, program marketing, or other program factor; program influence score is based on an allocation of 100 points between program and non-program influences; the no-program score is the likelihood the customer would have installed the same exact equipment in absence of the program.

[72] IL Policy Manual, Section 7.3 Free Ridership and Spillover.

Although the policy manual includes spillover and the TRM includes prescriptive methods for estimating participant and nonparticipant spillover, experts stated that there are areas of opportunity for improving the estimates of market effects and spillover. One expert noted it is “hard to measure nonparticipant spillover” and that Illinois is “consistently under-estimating non-part spillover,” which they described as “huge and growing and having influence that isn’t able to capture.” Another expert said that there is a reasonable amount of long-term market transformation, customer awareness, and natural adoption of EE due to programs, and utilities “should get credit for it.” One utility has started to split spillover into separate studies,[73] conducting an overarching sector-wide spillover survey.

[73] For more information on the evaluation plan, see http://ilsagfiles.org/SAG_files/Meeting_Materials/2014/June_10_2014_TAC_Teleconference/IL_Spillover_Research_FINAL_2014-06-06.pdf

3. JURISDICTIONAL SCAN

This section provides a high-level review of NTG policies across North America, including the use of gross versus net values, components of NTG values, and the use of prospective versus retrospective values. In addition, the Navigant team examined how overall EE policies and regulatory considerations are related to the choice of NTG policies and approaches, as well as how they may link to other policy mechanisms such as lost revenue recovery, risk-reward mechanisms, and integrated resource planning.

This effort leverages prior work conducted for an EE stakeholder group that included utilities and regulators in Iowa in 2015.[74] To update and refresh the 2015 data, the Navigant team contacted industry experts that conduct NTG work across all US states (reaching out to over 30 experts) to confirm the status of NTG policies in each state and indicate any changes that have taken place since 2015. This outreach was particularly helpful in providing insight and understanding into many of the NTG nuances and complexities that often are not fully provided in public documents. For example, the terminology and definitions specific to a given jurisdiction can result in policies that appear similar across jurisdictions but, in practice, can be quite different.

3.1 Descriptions of High-Level NTG Policies

This section describes the framework of the high-level review. The findings are organized into five general policy categories to provide a basis for comparisons across jurisdictions. The five policy categories used in this analysis are:

1. ***Overall NTG policy.*** This shows whether or not PAs must report savings and assessment against goals (including risk-reward mechanisms) at the gross or net level. Note that states that use a NTG of 1.0 are assumed to effectively be gross states because there are no upward or downward adjustments due to program attribution. Additionally, states that use common practice baselines are assumed to be gross savings states.
2. ***Definition of net savings (allowance for spillover).*** Within jurisdictions that use net savings for reporting, there is a wide variation of which aspects of NTG are allowed in terms of savings claims. Some states consider net savings to be net of free ridership but not to include any aspects of spillover. Other states allow different aspects of spillover (i.e., participant and nonparticipant) to be counted as achieved net savings.
3. ***NTG methods.*** Certain states (including the case study states above) have developed specific NTG methods for use in their states. This can include specifics regarding the calculations (e.g., survey batteries and analysis algorithms for self-report approaches). These methods (or protocols) may be recommended (allowing for certain adjustments each year), but certain states may adopt a more rigid approach (e.g., Pennsylvania) where the developed protocols are required to be used more consistently over time for similar programs.
4. ***Fixed or researched net savings.*** A number of net savings states lock in a fixed NTG value that applies to all, or at least most, programs. Note that while this has a prospective aspect to it (in that NTG is fixed prior to the program year), this is considered different than the team's definition of prospective NTG (below), which is typically based on researched values that can vary by program and measure.

[74] See: Navigant (2015). *Iowa Energy-Efficiency Net-to-Gross Report.* Prepared for the State of Iowa Department of Commerce Utilities Board. Link: https://efs.iowa.gov/cs/groups/external/documents/docket/mdax/mjax/~edisp/1201494.pdf

5. ***Prospective versus retrospective application of net savings values.*** Another practice is the prospective use of NTG values, whereby NTG values researched in a current program year are applied prospectively to future year(s) rather than retrospectively to the current or past program year(s). Once NTG values are established, they are essentially locked until an updated value is derived and applied prospectively for going-forward program design, setting targets, and incentives.

3.2 Jurisdiction Scan – Discussion

This section presents the results of the high-level review of NTG policies, with the discussion grouped into the policy categories set out above. Some overview findings include:

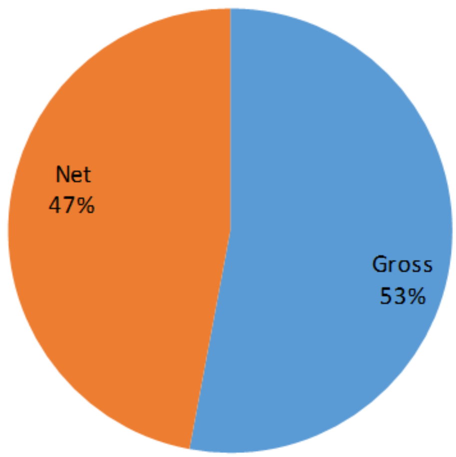
- Slightly over half of the jurisdictions (53%) use gross savings to assess whether energy savings goals and targets have been met. In some cases, gross savings are adjusted by in-field realization rate studies, and it is this adjusted gross value that is used to assess progress to goals and targets.
- Of those states that use net savings as the primary metric, nearly all of them (88%) include participant spillover and 67% include non-participant spillover.
- Most states that use net savings use either fixed NTG values or apply NTG on a prospective basis to facilitate program planning, progress to goals and targets, and use in determining shareholder incentives.
- There is a trend toward using a stakeholder process to develop agreed-upon methods for use in estimating NTG.
- States with financial incentives or other risk-reward policies are more likely to use net savings when assessing goals and incentives.

The balance of this section discusses the results of the jurisdictional scan for the five policy categories set out above.

3.2.1 Overall NTG Policy: Gross vs. Net

This policy category considers the role of gross and net savings in the reporting of energy savings and assessment against overall savings goals and targets. The research showed that slightly over half of the jurisdictions (53%) use gross savings in assessing achievement against savings goals. Note that gross states include states with limited (and in some cases minimal) utility efficiency programs. In addition, a number of states appear as gross savings states (e.g., Pennsylvania) but encourage and sometimes require NTG estimates and research to help improve program design. In fact, there is no state that the team is aware of that would prohibit NTG research for this purpose. This jurisdictional distribution of net versus gross is essentially the same as was found in the 2015 review of NTG policies.

Figure 1. Percentage of Jurisdiction Adoption of Net vs. Gross Policy

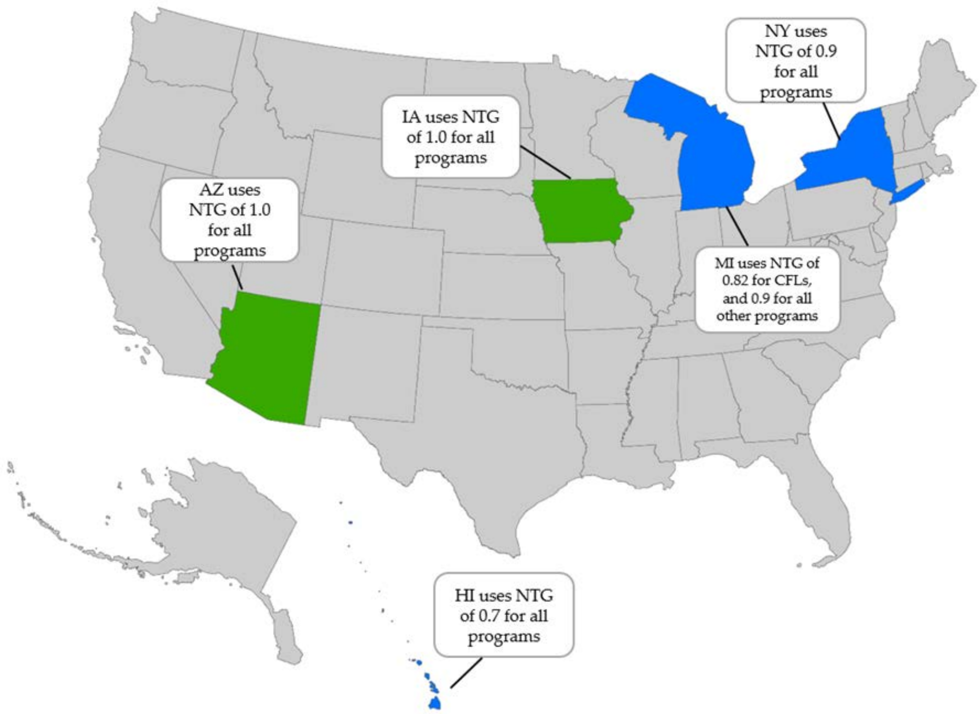


Source: Navigant team research

While some states adopt an overall net policy and others adopt a gross policy, a number of net states deem all program NTG values at 1.0 or a different value.[75] As shown in Figure 2, there are several states that currently adopt a NTG of 1.0 for all programs, while three states—Hawaii, Michigan, and New York—rely on fixed NTG values that differ from 1.0 for all programs within their EE portfolios. The ratios range from 0.7 to 0.9, and Michigan lowered the NTG for CFLs based on research showing a lower NTG than the other measures in the portfolio. Arkansas uses a similar approach during the first year of program implementation by having all programs use a stipulated NTG of 0.8. CFLs were the exception and were required to use an NTG of 0.62. After the first year, the programs were required to rely on researched values.

[75] Note this is considered different than prospectively deeming NTG ratios by program/measure based on researched values. These are global, portfolio-wide deemed values that are not regularly updated based on ongoing research. In addition, as noted above, states that deem NTG as 1.0 are counted as gross states for purposes of determining gross versus net policy.

Figure 2. States with Deemed NTG Values



Source: Navigant team research

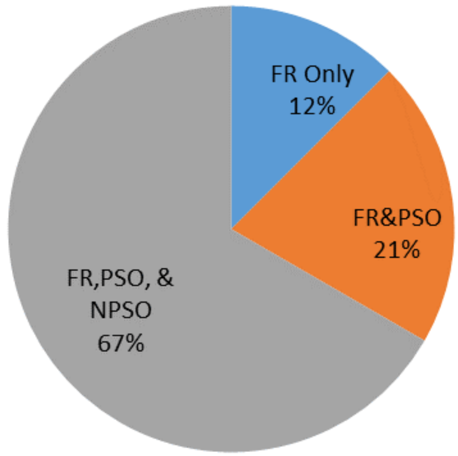
3.2.2 Definition of Net Savings (Allowance for Spillover)

As shown in Figure 3, over two-thirds (67%) of those jurisdictions that use net savings allow for free ridership, participant spillover, and nonparticipant spillover, while 21% allow for free ridership and participant spillover but do not allow for nonparticipant spillover.[76] Only 12% of the jurisdictions with net savings (a total of three states) limit net savings to net of free ridership (i.e., do not allow for contributions from spillover to count toward the net savings estimates).

---

[76] Note that the team is including market effects as a subset of non-participant spillover rather than breaking it out separately. This is because the team is not considering it in the context of market transformation studies, but rather as a subcategory of non-participant spillover. Precedent for this distinction is set by the US Environmental Protection Agency (EPA) in the *Model Energy Efficiency Program Impact Evaluation Guide: A resource of the National Action Plan for Energy Efficiency* (NAPEE), available at: http://www.epa.gov/cleanenergy/documents/suca/evaluation_guide.pdf

Figure 3. Inclusion of Free Ridership (FR), Participant Spillover (PSO), and Non-Participant Spillover (NPSO) in NTG Ratios (n=24)



Source: Navigant team research

3.2.3 NTG Methods Protocols

Several states reviewed as part of the jurisdictional scan have used open stakeholder processes to develop common methods for estimating NTG in their state. This can span the use of regression/statistical methods, but there often is a specific focus on developing agreed-upon methods for estimating NTG using self-report survey methods. States that have developed agreed-upon frameworks for survey methods include California, Massachusetts, Illinois, and Pennsylvania.

The primary reason for developing common approaches to NTG estimation is to help ensure that differences over time or between PA service territories are due to actual differences in program attribution as opposed to differences in research methodologies. In addition, bringing together PAs and their evaluators to develop these common methods is perceived as an opportunity to refine and improve existing methods, ensure transparency, and provide all stakeholders with greater confidence in the results of these studies. In addition, discussion and debate on the specifics of an applied method helps provide an appropriate policy and planning context around the resulting NTG estimates, which can be used to support good planning and policy decisions.

The primary drawback of common approaches, however, is that they can be perceived as inflexible and thus, inappropriate for programs that do not fit a typical model. They can also be seen as stifling innovation for new and potentially superior approaches. In response to these concerns, some of the more recent protocols have tried to retain some flexibility (e.g., giving example questions that can be adapted to specific program designs and features); plus, in some cases, they also allow evaluators to propose alternative methods that can attempted with regulatory approval.

The NTG methods guidance documents in both California and Massachusetts focus on the selection of methods (i.e., which methods are most appropriate for specific program types) as well details regarding the application of estimation methods for select programs. In general, more specific guidance has been developed for applying self-report methods, as these are often viewed as requiring greater judgment in their application. For example, these guidance documents can go as far as listing example questions and scoring algorithms used to determine NTG estimates. More recently, both Illinois[77] and Pennsylvania have followed the lead of California and Massachusetts by instituting common NTG approaches.[78]

Most portfolios offer such a diversity of programs and the guidance often focuses on the recommended NTG methods—i.e., the best methods for certain program types (e.g., downstream rebates). More detailed guidance on the application of NTG estimation methods is often developed for those programs that are viewed as more challenging in terms of producing NTG estimates or are viewed as more important to the overall portfolio's total savings. Recently, there has been an effort to expand the number of programs for which guidance in the application of methods are developed. For example, Pennsylvania and Illinois both include guidance for appliance recycling programs, and Illinois includes guidance for a common approach to estimate NTG for upstream lighting.

3.2.4 Prospective vs. Retrospective Applications

NTG ratios can be sensitive to the methods used, and any variation in NTG estimates directly results in a proportional change in the estimated net savings (i.e., a 1% drop in NTG results directly in a 1% drop in attributable savings). As a result, PAs (often gas or electric utilities) perceive significant risk and uncertainty with retrospective NTG application in terms of the actions needed to meet targets and calculate incentives. Thus, the majority of net states use NTG results prospectively rather than retrospectively. Prospective NTG application means that any updates to NTG values are applied in future program years, not in the year in which they are developed or to prior program years. As shown in Figure 4, over half of the jurisdictions with net savings (62%) use either a prospective/fixed NTG or a combination of prospective and retrospective.[79]

The prospective approach reduces risks for PAs; however, to be effectively implemented, it requires careful planning in terms of the timing and nature of the NTG research so that the results are still considered timely and applicable going forward. NTG research needs to be updated as markets and incentive structures change and as secondary research from other similar programs indicates increased variability in NTG values across programs.

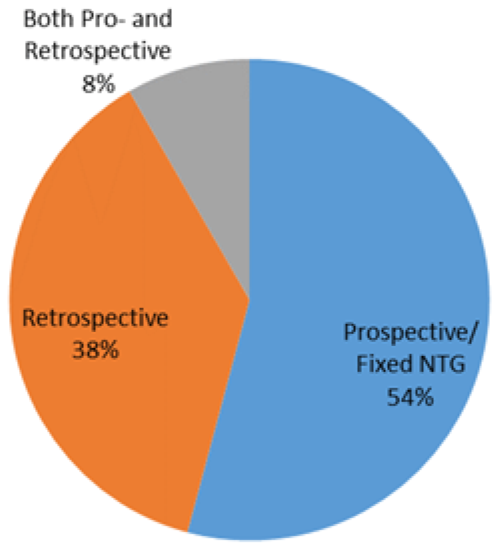
---

[77] IL TRM v6.0, Volume 4: http://ilsagfiles.org/SAG_files/Technical_Reference_Manual/Version_6/Final/IL-TRM_Effective_010118_v6.0_Vol_4_X-Cutting_Measures_and_Attach_020817_Final.pdf

[78] Jane Peters and Ryan Bliss, *Common Approach for Measuring Free-riders for Downstream Programs*, Prepared by Research Into Action as part of the Statewide Evaluation for the Pennsylvania Electric Distribution Companies and the Pennsylvania Technical Utility Staff, December 23, 2014.

[79] California, as part of the ESPI, allows prospective NTG for some measures but retrospective NTG for measures that are determined to be less stable in terms of program attribution.

Figure 4. Use of Prospective vs. Retrospective NTG (n=24)



Source: Navigant team research

3.2.5 Relationship between NTG Policy and Other DSM Policy Objectives

While a review and summary of NTG approaches, trends, and policies is helpful for supporting decision-making, it is also important to understand NTG policy in the context of other DSM policies. Certain jurisdictions may have DSM policies—particularly financial policies—where the importance of more precise, fully attributable savings estimations can be perceived as being more important. In addition, other DSM policies may be related to the treatment of NTG. The Navigant team examined a number of DSM policies and their relationships to NTG policy:

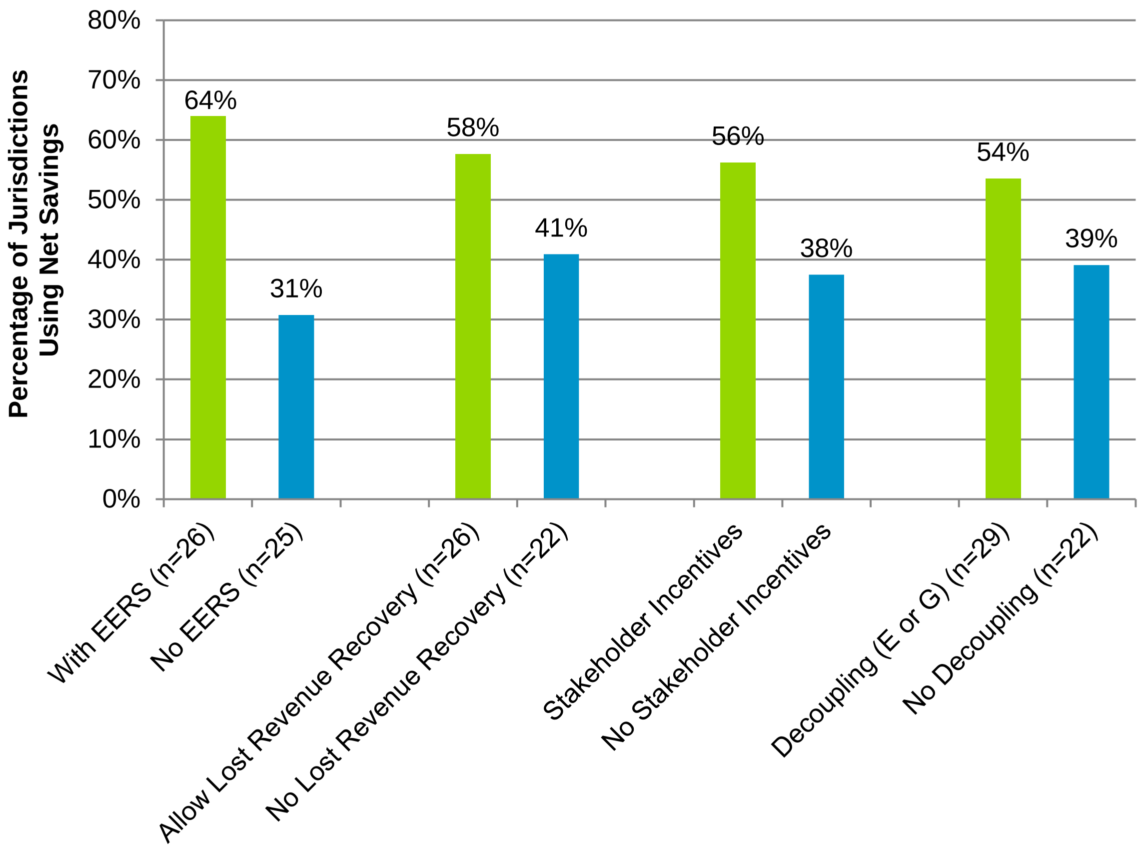
- **Energy efficiency resource standards (EERS):** State-level policy that sets long-term mandatory energy savings targets for utilities and EE PAs.
- **Decoupling:** A regulatory tool that serves as a means of helping utilities overcome the throughput incentive—i.e., the contribution to gross income that occurs with every energy unit sold because the unit (variable) price recovers some of a utility's fixed costs. A decoupling mechanism separates a utility's revenue from its unit sales volume without affecting the design of customer rates.
- **Lost revenue recovery:** Allows a utility to recover the lost revenue attributable to DSM programs by increasing revenue by that same amount. It can be based on decoupling (see above) or by adjustments (rate adjustment).
- **Risk-reward mechanisms:** Allows utilities to earn incentives for meeting or exceeding goals, or imposes financial penalties for savings shortfalls.

Figure 5 displays trends and correlations between the NTG policies and the broader DSM policy objectives. Note the sample sizes are small and each subset of analysis (e.g., examining only states that have net savings policies) leads to even smaller sample sizes. The results, therefore, need to be used with caution; they may not imply causation, but strong correlations do suggest that certain policies may be related.

The analysis focused on two of the most important NTG policies: the overarching policy of requiring savings goals and targets to be expressed as gross versus net savings, and for those jurisdictions that use net savings, whether or not the NTG is applied prospectively or retrospectively. These NTG policies are then compared against the different DSM policies outlined above.

As shown in Figure 5, jurisdictions that have an EERS allow lost revenue recovery and have bonuses or penalties that tend to also require net rather than gross savings. These differences are substantial; for example, 56% of the jurisdictions that have stakeholder incentives/penalties also require net savings, as compared to only 38% of those that do not have stakeholder incentives/penalties requiring net savings. Similarly, 58% of the jurisdictions that allow for lost revenue recovery require net savings, compared to only 41% of the jurisdictions that do not allow for lost revenue recovery. Each of these DSM policies with greater correlation with net savings also tend to be associated with increased stakes—particularly financial—in terms of the outcome. In other words, states where there are potentially millions of dollars on the line on the outcome of the savings assessment also tend to use net rather than gross savings.

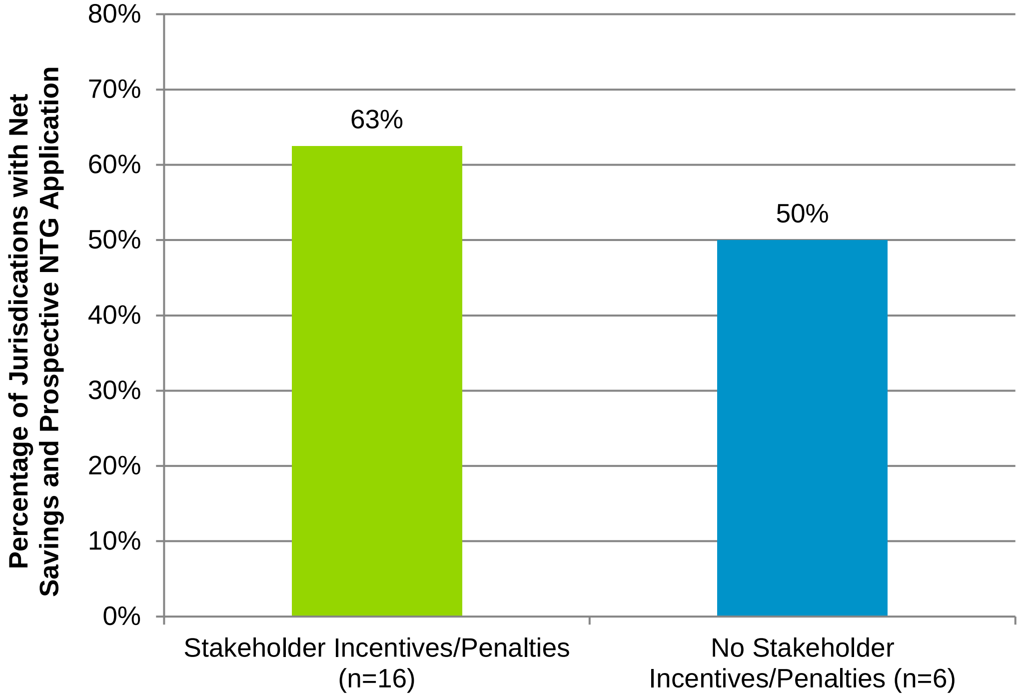
Figure 5. Percentage of Jurisdictions with Net Savings and a DSM Policy



Note, for example, the first green bar indicates that 64% of the states that have an EERS require net savings and the second green bar says that 58% of the states that allow for lost revenue recovery also require net savings, and the blue bar indicates the percentage of jurisdictions that do not have those policies that also require net savings.
Source: Navigant team research.

The Navigant team also examined the relationship of prospective versus retrospective NTG application and other DSM policy objectives. Due to the small sample size, the analysis was limited to only stakeholder incentives/penalties. As shown in Figure 6, jurisdictions with incentives are more likely (63%) to have prospective application of NTG findings versus areas without incentives (50%). Note that only six states with net savings did not have a bonus or penalty, so the sample size is quite small.

Figure 6. Percentage of Jurisdictions with Prospective NTG Application by Other DSM Policy (States with Net Savings Only) [80]



Source: Navigant team research

[80] Note that because California and Colorado have both prospective and retrospective NTG they are not included in this chart.

APPENDIX A. JURISDICTION NTG POLICY SUMMARY

	Use of Gross or Net Savings			Components of NTG				DSM/Cost Recovery Policies				
State	(A) Gross (savings assume NTG of 1.0) or Net?	(B) NTG Fixed at single value other than 1.0 for all measures?	(C) If (b)=Y then ask: What is the alternative value?	Participant SO	Non-Participant SO	Free Ridership	Goal Assessment: Gross, Fixed Net, or Adjusted (retro) Net?	Energy Efficiency Resource Standard	IRP Process	Decoupling	Lost Revenue Recovery	Risk / Reward
Alabama	Gross	No	N/A	N	N	N	Gross	No	No	No	Yes	Bonus
Alaska	Gross	No	N/A	N	N	N	Gross	No	No	No	No	No
Arizona	Gross assuming NTG of 1.0	No	N/A	N	N	N	Gross	Yes	Yes	E&G	Yes	Bonus
Arkansas	Net	No	N/A	Y	Y	Y	Retrospective	Yes	Yes	E&G	Yes	Bonus
California	Net	No	N/A	Y	N	Y	Fixed and Retrospective	Yes	Filing rqmt long-term plans	E&G	No	Both
Colorado	Net	No	N/A	Y	Y	Y	Fixed and Retrospective	Yes	Yes	No	Yes	Bonus
Connecticut	Net	No	N/A	Y	Y	Y	Fixed	Yes	Filing rqmt long-term plans	E, Pending G	Yes	Bonus
Delaware	Gross	No	N/A	N	N	N	Gross	No	Yes	Pending G	Pending	No
District of Columbia	Gross	No	N/A	N	N	N	Gross	No	No	E	No	Bonus
Florida	Net	No	N/A	Y	Y	Y	Retrospective	No	Filing rqmt long-term plans	No	Yes	Bonus
Georgia	Net	No	N/A	Y	Y	Y	Fixed	No	Yes	G	No	Bonus
Hawaii	Net	Yes	0.7	N	N	Y	Fixed	Yes	Yes	No	No	Bonus
Idaho	Gross	No	N/A	N	N	N	Gross	No	Yes	E	No	No
Illinois	Net	No	N/A	Y	Y	Y	Fixed	Yes	Filing rqmt long-term plans	G	No	Bonus
Indiana	Gross	No	N/A	Y	Y	Y	Gross	Yes	Developin g IRP	G	Yes	Bonus
Iowa	Gross assuming	No	N/A	Y	Y	Y	Gross	Yes	No	No	No	No

	Use of Gross or Net Savings			Components of NTG				DSM/Cost Recovery Policies				
State	(A) Gross (savings assume NTG of 1.0) or Net?	(B) NTG Fixed at single value other than 1.0 for all measures?	(C) If (b)=Y then ask: What is the alternative value?	Participant SO	Non-Participant SO	Free Ridership	Goal Assessment: Gross, Fixed Net, or Adjusted (retro) Net?	Energy Efficiency Resource Standard	IRP Process	Decoupling	Lost Revenue Recovery	Risk / Reward
	NTG of 1.0											
Kansas	Net	No	N/A	N	N	Y	Retrospective	No	No	No	Yes	Bonus
Kentucky	Gross	No	N/A	N	N	N	Gross	No	Yes	No	Yes	Bonus
Louisiana	Gross	No	N/A	N	N	N	Gross	No	Yes	No	Yes	Bonus
Maine	Net	No	N/A	Y	N	Y	Fixed	Yes	No	No	No	No
Maryland	Gross	No	N/A	N	N	N	Gross	Yes	No	E&G	Yes	No
Massachusetts	Net	No	N/A	Y	Y	Y	Fixed	Yes	Filing rqmt long-term plans	E&G	Yes	Bonus
Michigan	Net	Yes	0.9 for all measures except CFLs/LEDs, which are 0.82	Y	Y	Y	Fixed	Yes	Filing rqmt long-term plans	E&G	Yes	Bonus
Minnesota	Gross	No	N/A	N	N	N	Gross	Yes	Yes	G, Pending E	No	Bonus
Mississippi	Gross	No	N/A	N	N	N	Gross	No	No	No	Pending	Pending
Missouri	Net	No	N/A	Y	Y	Y	Fixed	No	Yes	No	Yes	Bonus
Montana	Gross	No	N/A	N	N	N	Gross	No	Yes	No	Yes	Pending
Nebraska	Gross	No	N/A	N	N	N	Gross	No	Yes	Pending G	No	No
Nevada	Net	No	Varies	Y	Y	Y	Fixed	No	Yes	G	No	No
New Hampshire	Gross	No	N/A	N	N	N	Gross	No	Yes	No	Yes	Bonus
New Jersey	Gross	No	N/A	N	N	N	Gross	No	No	G	No	No
New Mexico	Net	No	N/A	N	N	Y	Retrospective	Yes	Yes	No	Yes	Bonus
New York	Net	Yes	0.9	Y	Y	Y	Fixed	Yes	No	E&G	No	Bonus
North Carolina	Net	No	N/A	Y	Y	Y	Retrospective	Yes	Yes	G	Yes	Bonus
North Dakota	Gross	No	N/A	N	N	N	Gross	No	Yes	No	No	No
Ohio	Gross	No	N/A	N	N	N	Gross	Yes	Filing rqmt long-term plans	E	Yes	Bonus
Oklahoma	Gross	No	N/A	N	N	N	Gross	No	Yes	No	Yes	Bonus
Oregon	Net	No	N/A	Y	Y	Y	Fixed	Yes	Yes	E&G	Yes	No
Pennsylvania	Gross	No	N/A	Y	Y	Y	Gross	Yes	No	No	No	Penalty
Rhode Island	Net	No	N/A	Y	Y	Y	Fixed	Yes	Yes	E&G	Yes	Bonus
South Carolina	Gross	No	N/A	N	N	N	Gross	No	Yes	No	Yes	Bonus
South Dakota	Net	No	N/A	Y	N	Y	Retrospective	No	Yes	No	Yes	Bonus

	Use of Gross or Net Savings			Components of NTG				DSM/Cost Recovery Policies				
State	(A) Gross (savings assume NTG of 1.0) or Net?	(B) NTG Fixed at single value other than 1.0 for all measures?	(C) If (b)=Y then ask: What is the alternative value?	Participant SO	Non-Participant SO	Free Ridership	Goal Assessment: Gross, Fixed Net, or Adjusted (retro) Net?	Energy Efficiency Resource Standard	IRP Process	Decoupling	Lost Revenue Recovery	Risk / Reward
Tennessee	Gross	No	N/A	N	N	N	Gross	No	No	G	No	No
Texas	Gross	No	N/A	N	N	N	Gross	Yes	Filing rqmt long-term plans	No	No	Bonus
Utah	Net	No	N/A	Y	Y	Y	Retrospective	No	Yes	G	Yes	No
Vermont	Net	No	Varies	Y	Y	Y	Fixed	Yes	Yes	E	No	Bonus
Virginia	Gross	No	N/A	N	N	N	Gross	No	Yes	G	Pending	No
Washington	Gross	No	N/A	Y	Y	N	Gross	Yes	Yes	Yes	No	Incentive
West Virginia	Gross	No	N/A	N	N	N	Gross	No	No	No	No	Pending
Wisconsin	Net	No	N/A	Y	N	Y	Retrospective	Yes	No	No	No	No
Wyoming	Net	No	N/A	Y	N	Y	Retrospective	No	Yes	G	Yes	No

NAVIGANT

Memorandum: Discussion of Selected NTG Estimation Issues

Prepared for:

Enbridge Gas Distribution, Inc.

and

Union Gas Limited

Prepared by:
Navigant Consulting, Inc.
1375 Walnut Street
Suite 200
Boulder, CO 80302
303.728.2500

and

Apex Analytics, LLC
2500 30th Street, Suite 207
Boulder, CO 80302
303.590.9888

December 14, 2017

The Navigant team submitted the following memorandum to Deborah Bullock (Enbridge Gas) and Leslie Kulperger (Union Gas) discussing select NTG estimation issues. The memorandum covers the following:

NAVIGANT

Memorandum

To: Deborah Bullock (Enbridge Gas) and Leslie Kulperger (Union Gas)

From: Dan Violette (Navigant) and Scott Dimetrosky (Apex Analytics)

Date: December 13, 2017

Re: Discussion of Selected NTG Estimation Issues

This memorandum represents the Navigant team's perspective on a set of issues beyond the information gathered for the Jurisdictional Review report.[1] A discussion of six issues was requested:

1. Net-to-gross (NTG) scoring process used in the DNV *Custom Savings Verification and Free-ridership Evaluation* report prepared for the Ontario Energy Board (dated October 12, 2017)
2. Context around the use of NTG estimates given challenges in estimation
3. Best practices in the use of self-report survey methods
4. Addressing attribution when there are multiple programs operated concurrently by different entities (e.g., gas and electric utilities)
5. Baseline issues and their relationship with net savings estimations
6. Tailoring attribution studies to specific programs to appropriately address program objectives and delivery considerations

Each of these issues are addressed in the following sections.

Issue 1: NTG Scoring Process

The scoring process refers to the approaches used to translate survey responses into NTG values. Each survey respondent has an individual NTG value developed for them that is consistent with their responses to a set of questions. This makes it important to ask questions that respondents can understand and respond to with some degree of accuracy. Overly difficult questions or questions that require a level of precision in their answers that cannot reasonably be provided by program participants complicates the translation of survey responses into NTG estimates. Judgment is required in both developing the question batteries and the algorithms used to translate responses into NTG values.

The scoring algorithm is central to any resulting NTG estimates. As a result, it is important that the algorithms be as transparent as possible and undergo a stakeholder review process to build confidence in the approach. A process that allows for discussion of the scoring algorithms, includes sensitivity analyses to assess robustness, and is as transparent as possible is important for producing NTG values that will have buy-in from stakeholders.

[1] Navigant Consulting, Inc., *Net-to-Gross Policies: Cross-Cutting Jurisdictional Review*. Submitted to Enbridge Gas and Union Gas, December 14, 2017.

1.1 Ontario Free Ridership Evaluation Questions

DNV uses a methodology termed Life Cycle Net Savings (LCNS) to determine the free ridership component of NTG.[2] DNV indicates that "the treatment of timing is how LCNS differs from other estimation approaches for attribution." This overall approach uses three attribution parameters:

1. **Timing:** Did the program accelerate implementation of a measure or cause it to be implemented before it would have been without the program?
2. **Efficiency:** Did the program increase the efficiency of a measure above what would have been installed in the absence of the program?
3. **Quantity:** Did the program increase the quantity of a measure above what would have been installed in the absence of the program?

It is common for scoring algorithms to use more than one attribution parameter or influence factor. The method used by DNV focuses on program influence factors, but some algorithms will also include some non-program factors such as previous experience with the measure and organizational policy or guidelines.

This section focuses on the questions used to address the timing influence parameter in the DNV for brevity; however, a review of question wording and design for most any attribution parameter would have similar issues and considerations.

Timing – Questions Review

It is useful for reviewers to consider the questions asked in the telephone survey and their form. To assess the timing attribution parameter, DNV asked four questions of participants, as shown below:

LEAD IN -- Now I would like to get into some specifics of the <project_n>. I would first like to know about what effect, *if any*, that the <utility> <program> program had on your decision to perform the actions in that project *when you did.*

DTA1a Without the <utility> <program> program, would you have performed the <project_n > at the …

1	**Same time**	DAT1a
2	**Earlier**	
3	**Later**	
4	**Or Never?**	
98	Don't Know	
99	Refused	

[2] This is described in the *2015 Natural Gas Demand Side Management Custom Savings Verification and Free-ridership Evaluation* report prepared by DNV GL for the Ontario Energy Board. October 12, 2017.

DAT1a_O. Why do you say that? [RECORD VERBATIM]

77	Record Response	IF DAT1a = NEVER, SKIP TO DAT1c] [ELSE IF DAT1a ≠ LATER, SKIP TO DAT2a]
98	Don't Know	
99	Refused	

DAT1b. Approximately how many months later?
[Try to get a number. Try bracketing if necessary by beginning with more or less than four years later.]

1	Record Number of months	
98	Don't Know	
99	Refused	

DAT1c. How old was that equipment?
[Get age at time of replacement. If they cannot provide exact age, ask for year installed and calculate age.]

1	Record Age	DAT2a
98	Don't Know	
99	Refused	

These questions have a logical flow, and the question *DAT1a_O* about why a respondent selected the answer to the initial question to be "same time, earlier, later or never" provides additional insight into the participant's response. DNV provided the full set of verbatim responses, which helps in transparency and in assessing whether the questions were understood by the respondents.

This bank of questions is aligned with approaches used in attribution studies across the industry, and there is no single right way to ask these questions. Yet, there are some judgments made in the way questions are structured and how the responses are used in the scoring algorithm. As a result, it can be important for stakeholders to review the underlying judgments and assumptions to build confidence in the overall process.

Timing – Judgement in Questions and Scoring

One example of judgment in the survey question development and scoring is the DNV 48-month cutoff for partial free ridership. To illustrate, if the answer to question *DAT1a* is "later," then question *DAT2b* asks the respondent how many months later they would have performed the project. This is meant to get an estimate of when—if the program had not existed—the participant would have performed or installed the project. The person conducting the survey is directed to "try to get a number" and, if helpful, try bracketing by asking if the project would have been performed more than 4 years later (i.e., 48 months) or less than 4 years later. An approach like this can be used to help the respondent begin to think through their answer. Also, in the algorithm used by DNV, any project that would have been undertaken more than 4 years later is given an NTG value of 100% (i.e., zero free ridership). For estimates of less than 48 months, a partial attribution credit is developed, which produces an NTG less than 100%.

The use of 48 months as the cutoff is an informed judgment made by the DNV evaluation team. The importance of this assumption can be addressed through sensitivity analysis. For example, alternative cutoff values could be considered. For example, if respondents are saying the project would be delayed by more than 3 years (i.e., 36 months), would it be reasonable to assume that they are not free riders, as this may be a speculative response? Using 36 months as the cutoff in the scoring algorithm instead of 48 months would provide information to stakeholders about the importance of this selected cutoff value.

Timing – Question Design Review

A review of the timing questions involves thinking though how well participants are able to provide these values. DNV states that timing was the most important attribution factor for many of the programs evaluated. With respect to Union's Custom C&I program, DNV states, "Timing was the component most strongly affected by the program. The program affected the timing of projects that account for approximately half of the energy saving" (p.35). Given this context, it is important to consider whether respondents can provide accurate estimates of when they would have performed their recent project if the utility program had not existed and if this information can be provided at a monthly level. Is it reasonable to distinguish between 6 months, 9 months, or 12 months later?

Looking at some of the verbatim responses can give a reviewer a sense of how accurately respondents might be able to assign the month in which they would have performed the project undertaken as part of the utility-offered program. A few of the verbatim responses from participants that said they would have performed the project at a later date are shown below (from Table 8-9: Timing Verbatim Responses Union Custom C&I programs, p.C-2):

- At some point in time we would have learned the value of this and done it.
- It's one of those things that you put on a list and OK, we'll do it sometime, but it might be 5 years or 3 years. Hard to say.
- May have done it the next year without incentives. Hard to say if upper management would have approved.
- Probably would never have done it; if so, maybe a couple of years.
- They will wait to replace something until they really have to, unless it's a health and safety issue.
- It would have taken longer to get approval.
- Probably would never have done it; if so, maybe a couple of years.
- They would've had to do these eventually.
- Tough question - It's possible that we just would have done nothing at all. Maybe fewer if we did.

These verbatim responses would seem to imply that it might be difficult for these respondents to determine the month in which they would have performed the project implemented under the utility program if the program had not been offered. This is not a criticism of the DNV approach, but it does show how stakeholder review may be helpful in assessing question form, the use of the monthly time period, and the resulting NTG value. For additional discussion of the effect of this response on the NTG result, see Section 1.2.

Timing – Potential for Response Bias

One concern when questioning a program participant about whether they would have undertaken the energy efficiency (EE) investment even if the program had not been offered is the social desirability bias: yes, they would have undertaken the EE investment in the absence of the program, as they may want to view themselves as a good citizen with an understanding of the benefits of EE investments. This translates into concern that respondents may be overly confident about how quickly they would have undertaken the project in the absence of the program.[3] As a result, statements about future respondent EE actions in the absence of the program may be biased high and the NTG from this set of timing questions may be biased low.

Given this concern, it might be useful to conduct sensitivity analysis to test the difference in the NTG values that would result if participant estimates were overly optimistic. One scenario might be that participants are overly confident in the actions they would have taken by 50%. In this scenario, a sensitivity analysis would be conducted by changing responses of 3 months later to 6 months later, and if the response was 6 months, the sensitivity would use 12 months in the NTG calculation. This analysis would provide information on how sensitive the overall NTG values are to response errors in the timing question. This would allow stakeholders to comment on whether is it reasonable to consider an outcome where the respondents were overly confident about their undertaking the project in the absence of a program.

Reviewing potential response bias and the role of judgment in the survey design and scoring algorithm is not meant to be a critique of DNV's approach, as developing and using counterfactual "what if" questions are difficult in all EE attribution studies. Additionally, this problem is not unique to EE evaluation. These counterfactual scenarios must be addressed in any evaluation performed across a wide number of fields including the evaluation of business management decisions (e.g., the benefits and costs of offering a benefits package to employees), health programs (e.g., a school lunch program), and any other evaluations of policies and investments.

Timing – Early Retirement

The fourth question in the battery of timing questions asks about the age of existing equipment, which may also be important in the NTG scoring as part of DNV's LCNS methodology. The question on equipment age is:

[3] The New York State Department of Public Service, *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approach* states, "Often a series of survey questions are asked of the participant about the actions they would have taken if there had been no program to derive a free ridership estimate. More specifically, this is asking the respondent to state their intentions with respect to purchasing the relevant equipment absent the program. Bias creeps in because people may intend many things that they do not eventually accomplish." Link: https://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

DAT1c. How old was that equipment?
[Get age at time of replacement. If they cannot provide exact age, ask for year installed and calculate age.]

The response to question *DAT1c* above is used in conjunction with an estimate of the remaining useful life (RUL) of equipment replaced to allow for the use of different efficiency baselines for equipment that is viewed as being replaced early versus equipment installed on failure. This is a component of DNV's LCNS methodology, which is presented in detail in Appendix J of the DNV report. In general terms, these baselines are:

- If it is an early replacement, the baseline for the years of RUL is the current efficiency levels of the replaced equipment. After that, the efficiency baseline is the industry standard efficiency for new equipment that is installed at that time.
- If the equipment is replaced on failure, then the industry standard efficiency is used for the expected life of the equipment.

These calculations require information on the age of the equipment replaced by the EE project (as provided by the respondents), estimates on the useful life of the equipment, and the expected life of the new equipment installed. Can the respondents provide the age of the equipment or the installation date of replaced equipment? This is probably true, but there may still be uncertainty and assumptions in this lifecycle savings estimate.[4]

1.2 Comparison of Scoring and Attribution between Ontario and Massachusetts

The questions used in the Ontario Free-ridership Evaluation are very similar to those used in Massachusetts.[5] The similarity of the questions allows for a comparison of calculated NTG values for the same responses to specific questions across the two jurisdictions.

DNV provides examples of attribution calculations that can be used in this comparison. Table 8-86 in the DNV report provides several examples of how survey responses are translated into an NTG ratio. The first row of Table 8-86 (reproduced below) states that the effect of project "acceleration only" with a participant response indicating that they would have undertaken the project 24 months later if the utility program had not been offered results in an NTG of 31%.
If the "months later" response was 48 months, DNV's algorithm assigns an NTG of 100% (i.e., free ridership is zero), as this is the cutoff value. In this example attribution calculation in Table 8-86, the "months later" is 24 months, a value that is one-half of the value that would produce an NTG of 100%. These results are summarized in the table below:

Months Later Response	NTG Score	Free Rider Score
24 months later	31%	69%

[4] Additionally, it is worth noting that the DNV approach applies early replacement credit only to those participants who cite the program for accelerating the timing of their investment. In some jurisdictions, such as California, this type of adjustment is handled in the gross savings realization rate estimation and does not require the program to influence project timing in order to claim early replacement credit[4]. Therefore, when net savings factors account for early replacement, it is necessary to ensure that double counting does not occur when estimating the gross savings realization rates. Although we assume this did not happen in the DNV study, Appendix J is not explicit in whether early replacement savings adjustments are accounted for in gross realization rates and if so, how DNV ensures that double counting of the gross realization rates and NTG adjustments is not occurring.
[5] This is not unexpected at DNV is working on the customer C&I NTG studies in both Ontario and Massachusetts.

48 months later	100%	0%

The table above shows that a "24 months later" response produces an NTG value of 31% (e.g., a free rider value of 69%). This is a non-linear relationship—the "24 months later" response is halfway to the "48 months later" response, but the difference in NTG scores is much greater (100% compared to 31%). Additionally, the free ridership scores differ quite a bit as well – 0% to 69%. The reason why this is not a linear relationship is likely due to the way DNV addresses early replacement within its LCNS methodology. However, there is not enough information presented in the report to replicate this value.[6]

[6] It could be the case that DNV presented these calculations in workshops in Ontario, as the Navigant team was not involved throughout the entire NTG stakeholder process. It is not uncommon for these calculations to be presented in greater detail in workshops, but the report does not contain the information necessary to track how this example attribution was made.

Table 8-86. Ontario Attribution Examples[7]

Example	Timing Response		Efficiency Response		Quantity Response		VGS_E	VGS_S	$Y_{V.RUL}$	$Y_{V.EUL}$	VGS_L	Y_A	E	Q	SPA	NS_L	NTG
Accl only	Later	Two Years	Same		Same		100	50	3	10	650	2	0%	0%	0%	200	**31%**
"Never" for timing	Never		Same		Same		100	50	3	10	650	3	0%	0%	100%	650	**100%**
No attribution	Same		Same		Same		100	50	3	10	650	0	0%	0%	0%	0	**0%**
Accl with partial efficiency	Later	Two Years	Less	Between	Same		100	50	3	10	650	2	50%	0%	50%	400	**62%**
"Never" with partial eff.	Never		Less	Between	Same		100	50	3	10	650	3	50%	0%	100%	650	**100%**
Partial eff. only	Same		Less	Between	Same		100	50	3	10	650	0	50%	0%	50%	250	**38%**
Accl with partial eff. and partial quantity	Later	Two Years	Less	Between	Less	Half	100	50	3	10	650	2	50%	50%	75%	500	**77%**
"Never" with partial eff. and partial quantity	Never		Less	Between	Less	Half	100	50	3	10	650	3	50%	50%	100%	650	**100%**
Partial efficiency and partial quantity	Same		Less	Between	Less	Half	100	50	3	10	650	0	50%	50%	75%	375	**58%**
"None" is equal to "Never"	Same		Same		None		100	50	3	10	650	3	100%	0%	100%	650	**100%**
Full eff. credit, no accel. or quantity (ER)	Same		Less	Standard	Same		100	50	3	10	650	0	0%	100%	100%	500	**77%**
Full eff. credit, no accel. or quantity (non-ER)	Same		Less	Standard	Same		0	50	0	10	500	0	0%	100%	100%	500	**100%**

[7] Appendix K, Table 8-86, *2015 Natural Gas Demand Side Management Custom Savings Verification and Free-ridership Evaluation*. Prepared by DNV GL for the Ontario Energy Board. October 12, 2017

Because the same questions are asked in Massachusetts, it is possible to compare NTG and free rider scores for the same responses. For the same 24 months later response, the resulting NTG value is higher in Massachusetts—i.e., a 50% NTG in Massachusetts compared to 31% in Ontario. This a 60% increase in the NTG value (i.e., 50% NTG/31% NTG) due to different scoring algorithms. The reasons for this difference cannot be diagnosed in this review due to incomplete information, and there are no judgments about one NTG score being more appropriate than another NTG score; however, it is something that stakeholders in both jurisdictions might want to better understand.

A second attribution example calculation from Table 8-86 using the efficiency attribution parameter also demonstrates the role of judgment in the NTG scoring algorithm. Table 8-86 indicates that a program participant that states they were influenced to:

1. Install the high efficiency unit through the utility program, but
2. In the absence of the program, they would have installed a unit that is:
    a. Higher than a standard efficiency unit; but,
    b. Lower than the efficiency of the program unit they installed.

These participants also state that they would have installed the equipment at the same time and in the same quantity. These responses put them in the partial free rider classification.[8]

Given the responses set out above, the NTG value assigned by the DNV algorithm in Ontario is 38% (from Table 8-86). The same responses in Massachusetts would produce an NTG value of 50%. In terms of free ridership, Massachusetts would have a free ridership score of 50% compared to a 62% free ridership score in Ontario. This is another example of how differences in the scoring algorithms can influence NTG values, even if the questions are very similar.

As a note, there are additional factors accounted for in Massachusetts that influence the NTG scores. The same question on efficiency partial free ridership is currently included in the Massachusetts algorithm; however, Massachusetts is considering removing this question because including an intermediate efficiency response in the NTG algorithm and also through the use of industry standard practice (ISP) baselines in estimating gross savings could result in double counting the efficiency penalty.[9]

Recent research in a forthcoming northeast US study has shown the responses to these partial free rider questions can be highly variable.[10] For example, when those respondents that selected the response option of “between standard efficiency and what you purchased” were further asked what they would have purchased, a number of respondents said they would have purchased equipment that met code; or, the least expensive option. This invalidates their prior response that they would have purchased intermediate efficiency levels. In addition, the most common response to these questions was “don’t know” or “we weren’t considering additional efficiency levels.” This shows a potential the lack of consistency in responses to this question on partial free ridership. It also demonstrates the value of using consistency check questions in the survey.

---

[8] As a note, there may be concerns about how respondents understand what the standard efficiency baseline represents in this question, particularly if they also respond that they would have taken this action 24 months later.

[9] See the *Massachusetts Commercial/Industrial Baseline Framework.* prepared by ERS and DNV GL for the Massachusetts Program Administrators, April 26, 2017 (http://ma-eeac.org/wordpress/wp-content/uploads/MA-Commercial-and-Industrial-Baseline-Framework-1.pdf)

[10] Tetra Tech. *2016 Commercial and Industrial Programs Free-Ridership and Spillover Study* (Draft), Prepared for National Grid Rhode Island, September 2017.

Another example of where judgment plays a role is in the assignment of NTG values to question responses concerning partial free ridership. When a participant responds with the answer that they would have purchased "between standard and efficient equipment" in absence of the program, the Ontario algorithm decreases the efficiency attribution by 50%.[11] However, the actual decrease in efficiency might be higher (e.g., 90%) or lower (e.g., 10%) depending on equipment specifications and the resulting savings relative to the efficiency of the program-installed unit. The effect of this 50% assumption could be examined through sensitivity analyses. In addition, surveys on a test sub-sample of participants can try to get more refined estimates that can be used to inform the scoring algorithm.

It should be noted that while there is evaluator judgement embedded in the Ontario methodology, this is also true for most all other methodologies for NTG scoring. The question for stakeholders is whether the expert evaluator judgment seem reasonable, and whether the NTG values are stable and robust across a reasonable range of assumed values.

1.3 California and Illinois NTG Scoring Examples

The California and Illinois NTG survey methods use similar NTG questions, but they are based on an alternative approach to the one used in Ontario and Massachusetts. These NTG values are based on three program attribution components, or indices:

1. **Program attribution index 1 score (PAI-1)** reflects the influence of the most important of the various program and program-related elements in the customer's decision to select the specific program measure at the time.
2. **Program attribution index 2 score (PAI-2)** captures the perceived importance of the program (whether rebate, recommendation, training, or other program intervention) relative to non-program factors in the decision to implement the specific measure that was eventually adopted or installed. The program influence score is adjusted (i.e., divided by 2) if respondents said that they had already made their decision to install the specific program-qualifying measure before they learned about the program.
3. **Program attribution index 3 score (PAI–3)** captures the likelihood of various actions the customer might have taken at the time and in the future if the program had not been available (the counterfactual).

The algorithms in California and Illinois are not very comparable to the one used in Ontario. However, in contrast to Ontario, the California and Illinois scoring algorithms take the maximum of the timing or efficiency scores for the PAI–3 score, meaning the program only has to influence either timing or efficiency to receive the full value of that score. This tends to produce higher NTG scores and lower free ridership scores. This free ridership score is also tempered by the PAI–1 and PAI–2 scores that address a wider range of influence factors (both program and non-program factors) that were influential in their decision-making process.

Issue 2: Context Around the Use of NTG Estimates

This discussion examines the statistical error in NTG estimates and its relationship to overall uncertainty in the NTG estimates. Statistical error addresses one aspect of the uncertainty in NTG estimates, but there are a number of other important contributing factors. A study can produce a point estimate of NTG with a high level of confidence and precision, but the overall uncertainty in the estimate may be far

[11] DNV GL. *2015 Natural Gas Demand Side Management Custom Savings Verification and Free-ridership Evaluation*. Prepared for the Ontario Energy Board. October 12, 2017. Appendix K, Table 8-84.

greater. Stakeholders should understand the uncertainty in the NTG estimates when they are considering how to these estimates for planning and potentially when calculating shareholder incentives.

2.1 Statistical Error in Survey Methods

The context of NTG estimates involves understanding the representations of confidence and precision accompanying these estimates. DNV has a good explanation in a section titled "Understanding Statistical Error" on page 18 of the report. DNV designed the samples used in the studies to target 10% relative precision with 90% confidence (90/10) based on the best available assumptions at the start of the evaluation.

These statistical criteria may seem too precise given the discussion of issues in scoring algorithms presented in Issue 1 above. Scoring algorithms may be based on questions that can be difficult for program participants to answer, and evaluator judgment is used in the construction of NTG scoring algorithms. The crux of the issue is that there is a difference between uncertainty[12] and statistical error.

Statistical error in the context of self-report survey findings represents a somewhat narrow concept. Usually a sample of participants is selected for the survey as it would be too expensive to survey the entire population of participants. Statistical error in this context refers to the relationship between the estimates and findings from the survey, and what would have been obtained if the entire population had participated in the survey instead of only a sample. A sample NTG estimate of 80% obtained from the survey with a confidence of 90% and a relative precision of 10% indicates that if the entire population was surveyed (to eliminate sample error), there would be a 90% probability that the value obtained from the entire population would fall between 72% and 88%, i.e., it would fall within an interval of 80% +/- 8%. In this case, the confidence and precision only represents sampling error and does not capture any issues related to response bias, or judgments applied by evaluators (even expert evaluators) in the construction of the NTG scoring algorithm.

The result is that attribution studies can produce NTG values that have 90% confidence and 10% precision, but addressing statistical error only may not appropriately dimension the overall uncertainty in the NTG values. This can be due to (real or perceived) biases in the inputs to the NTG scoring algorithm—due to biases in the responses to NTG questions and in accuracies in evaluator judgment used in the scoring algorithms. The effect of these factors can be assessed using sensitivity analyses (as discussed above) or other simulation methods that show the potential impacts of alternative inputs to the NTG algorithms. This is an important component of any assessment of an NTG scoring algorithm and should be part of a stakeholder review of the NTG estimation process.

2.2 Assessing Reliability and Consistency of NTG Estimates

Assessing the overall robustness and consistency of the NTG scoring algorithm across a range of reasonable assumptions and scenarios should be a part of any overall evaluation process. The result will typically provide a range for the NTG estimates. The case studies in the Jurisdictional Review report indicated that all three states used a collaborative, transparent stakeholder process to finalize NTG estimates.

This review process can highlight issues with the survey questions and format. One issue that often comes up is whether the survey provided enough context and explanation of the program to remind the

---

[12] Definitions of uncertainty in the sciences can be a complex concept; however, a definition for uncertainty that is useful in this context can be found on Wikipedia: "The lack of certainty, a state of limited knowledge where it is impossible to exactly describe the existing state, a future outcome, or more than one possible outcome."

participants responding to the survey of the program features, information, and potential influences. If a survey is conducted 1 year or more after participation in a program, the respondent may not recall all the features of the program and all the assistance provided. Instead, respondents may focus narrowly on the influence of the rebate or incentive payment.[13] It is also difficult for respondents to isolate the influence of one utility's programs from other influences if multiple programs are offered or if incentives are provided by more than one entity (e.g., incentives by a gas and an electric utility).

Sometimes a triangulation process is used where trade allies are surveyed to get their views on the influence of a program on selected participants or on the market as a whole. On occasion, trade allies can have different opinions on whether program participants would have really undertaken the actions offered through the program even if the program had not existed. Past industry experience has indicated that some participants may be overly optimistic regarding what they would have done on their own—i.e., in the absence of the program. Trade allies may be able to offer responses that are less prone to this bias and provide a means for adjustment.

The New York State DPS EM&V Guidelines states that "when multiple questions, weights, and complex algorithms are involved in calculating the NTG ratio, evaluators should also consider conducting a sensitivity analysis (e.g., changing weights, changing the questions used in estimating the NTG ratio, changing the probabilities assigned to different response categories, etc.) to assess the stability and possible bias of the estimated NTG ratio."[14] The NY DPS guidelines go on the state that "the onus is on the evaluator to demonstrate that the algorithm is not biased." However, the team would amend this to be the responsibilities of an overall stakeholder or evaluator review group in a jurisdiction. All parties have an interest in robust, stable results.

2.3 Promoting a Collaborative Process

An important component of NTG context depends on how the estimates are to be used. If the NTG estimates are to be used to calculate shareholder incentives, then the review and estimation process can become much more contentious. The potential for bias in responses to survey questions and from the survey frameworks (e.g., inadequate context on the overall program for respondents) can result in disagreements about exact NTG values or the fairness of the application of NTG estimates in an incentives calculation. The fact that some evaluator judgment is a part of every NTG scoring algorithm only compounds the discord.

In the Massachusetts case study,[15] the use of NTG retrospectively in incentive calculations was described as causing tension in the system, with significant disagreements over the NTG estimates. It was

---

[13] The New York State Department of Public Service, "*Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approach"* states, "This makes it essential that the interviewer guide the respondent through a process of establishing benchmarks against which to remember the events of interest. Failure to do so could well result in, among other things, the respondent 'telescoping' some events of interest to him into the period of interest to the evaluator. Set-up questions that set the mind of the respondent into the train of events that led to the installation, and that establish benchmarks, can minimize these problems." Link: https://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

[14] The New York State Department of Public Service, *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approach* p. 10 Link: https://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

[15] See Section 2 in Navigant Consulting, Inc., *Net-to-Gross Policies: Cross-Cutting Jurisdictional Review.* Submitted to Enbridge Gas and Union Gas, December 14, 2017.

described by one expert as “really, really bad in the past.” Another expert described NTG results to be “extremely negative as retrospective tool.” A third expert noted, with respect to incentives, that when program administrators (PAs)—typically utilities—are “losing money based on subjective studies, it gets ugly.” Therefore, in the most current cycle, Massachusetts adjusted its policies to apply agreed-upon NTG factors on a prospective basis. One best practice is to develop net savings and NTG approaches that are transparent, understood by the parties, and seek consensus where possible.

The case studies in the Jurisdictional Review contain examples where stakeholder groups have enhanced collaboration throughout the program planning, evaluation, and incentive calculation process. In all three states, stakeholders had the opportunity to question, challenge, and suggest modifications to the initial estimates produced by an EM&V study. One comment that was made was that because NTG “answers are not easy” there needs to be a process with “enough room for reasonable people to disagree.”

Another best practice is to use agreed-upon, pre-defined C&I NTG survey questions and algorithms. These are often viewed as common practice for that jurisdiction, but they are not strictly required or followed. This pre-defined approach allows NTG results to be compared over programs and time, and allows the algorithm to be updated, tweaked, and improved. This leads to a more sophisticated approach and increases stakeholder confidence, as it has been incrementally changing over time as part of a transparent review process.

Issue 3: Discussion of Best Practices

This section discusses some components of best practices in applying self-report surveys to C&I custom EE programs. This is a broad topic and too expansive to fully address in this memorandum. The Jurisdictional Review report, through the three state case studies, discusses a number of best practices. Given concerns with self-report methods, experts interviewed for that report noted multiple approaches that can be used improve the accuracy of self-report studies and thus promote confidence in the study findings. These same points are set out in the Jurisdictional Review report, but to provide the information in this document, key points are presented below:

1. **Fast feedback:** Fast feedback refers to survey methods where the respondents are asked about factors influencing their participation in a program at a time near to when they participated—e.g., within 3 months of completing participation. Experts noted the value in using fast feedback to gain the most accurate responses for free ridership, but it is not required in any state. A number of Illinois utilities use a parallel path evaluation approach for selected custom projects that allows for real-time NTG. In California, they are pre-screen custom projects with respect to an estimating initial NTG value to reduce risks of surprise NTG values when the full impact evaluation is performed. This two-step approach helps produce a “no surprise” approach that builds confidence in the NTG estimates.
2. **Sensitivity analysis:** Sensitivity analysis (with full transparency regarding the scoring) has been used in all three case study states, primarily when the pre-defined batteries are first developed and tested, but the algorithms are also periodically revisited. This can be important as different but reasonable assumptions used in translating question responses into NTG values can result in different NTG values.
3. **Triangulation:** The perspective of vendors is collected in all states for custom projects on a project-by-project basis (e.g., if the customer states the trade ally recommendation was important) and can increase the NTG result. Triangulation with vendors/trade ally surveys is also used to address the influence of factors that program participants may not be well positioned to address—e.g., the relative influence of multiple program influences on program impacts. As noted below, multiple experts noted the difficulty of participants understanding attribution of any

individual influence on their decision-making, as there are many potential influences in the EE marketplace.

4. **Other best practices*:*** Other best practices mentioned by experts include: including multiple factors in the NTG scoring (program influence and other non-program influences), ensuring the questions and weighting are fully vetted, consistency checking, and gaining insight into the project story by spending additional time with the participant to understand.

Information on best practices are contained in the references to documents for the three case study states (Massachusetts, California, and Illinois). Three other documents providing guidance on the use of self-report surveys for estimating net savings are:

1. Violette, Daniel M.; Rathbun, Pamela. *Chapter 21: Estimating Net Savings – Common Practices: Methods for Determining Energy-Efficiency Savings for Specific Measures. National Renewable Energy Laboratory*, 2017. NREL/SR-7A40-68578. https://www.nrel.gov/docs/fy17osti/68578.pdf *[cut and paste link into browser]*

2. New York State Department of Public Service, N*ew York Evaluation Plan Guidance for EEPS Program Administrators -- Appendix G- Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approach,* 2013. Updated August 2014. https://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

3. Research Into Action. *Review and Analysis of Net-to-Gross Assessment Issues for Natural Gas Demand Side Management Custom C&I Programs.* Prepared for Enbridge Gas Distribution, Inc. August 25, 2017.

Issue 4: Addressing Attribution with Multiple Programs

The landscape of utility program evaluation is made more complex when there are multiple programs targeting the same customers. The Jurisdictional Review report addressed the issue of overlapping programs. In that report, experts reported that it is difficult for program participants to disentangle the influence of multiple programs (e.g., when more than one entity is providing incentives or information to encourage program participation), and they recommend viewing simultaneous programs as a single offering for free ridership purposes. However, this may be less than satisfactory to stakeholders when a utility has shareholder incentives and the goal is to have those incentives tied directly net savings attributable only to their programs.[16]

If utility-specific net savings estimates are viewed as necessary then one approach might be the use of a stakeholder review process to finalize NTG estimates parsed out for each utility. In the Jurisdictional Review report, all three case study states indicated that stakeholders had the opportunity to question, challenge, and suggest modifications to the initial NTG estimates produced by the evaluations. There was a recognition that all NTG estimation methods face challenges in application, and working toward agreed-upon NTG values informed by an NTG evaluation was worth the effort.

Another issue when dealing with the influence of multiple programs is the stacking of incentives. This occurs when a utility can help a customer not only obtain incentives from its program but also can help

---

[16] Problems in parsing out the influence of multiple programs is also discussed in: Research Into Action (2017). *Review and Analysis of Net-to-Gross Assessment Issues for Natural Gas Demand Side Management Custom C&I Programs.* Prepared for Enbridge Gas Distribution, Inc. August.

the customer qualify for incentives offered by other entities. When this occurs, it is difficult to parse out the influence of the utility that may have first worked with the customer. For example, one verbatim response to the timing question on Enbridge's C&I program indicated that Enbridge made the customer aware of other incentives. Estimating the partial contribution of one incentive over another may be difficult. The respondent may simply put the greatest weight on the largest incentive. However, the utility that initially interacted with the customer may have provided the impetus for program participation. There is a concern that some customers, when responding to influence questions, remember only the financial incentives and not the engagement, informational, and process elements derived from working with a utility to assess EE investments.

The issue of multiple programs and influences was addressed in a recent update to the New York State DPS EM&V Guidance, which stated that:

> *… some level of net savings assessment, or examination of program influence, can serve as an effective tool for program design and implementation. However, given the variety of activities occurring in the marketplace, including Commission direction for NYSERDA and utility offerings to become more complementary in nature, it will be increasingly more difficult to parse out the effects of any one specific program action.*[17]

Multiple programs and influences is an issue for other jurisdictions and regions as well. One regional study estimated total net savings from all the programs in a C&I sector, then a trade ally/market expert survey was conducted to allocate these savings to specific programs. In this case, the trade allies were judged as having a better perspective on the relative influence of individual programs than the customers would, as those customers might not even know how all of the programs interacted.[18]

Trying to parse out the influence of a single program in a market with multiple programs, incentives, and market influences is increasingly becoming a challenge. As was recognized in the New York updated guidance, it remains important to perform net savings assessments, but the limitations of the methods need to be considered. Good decisions do not require perfect information; rather, they require information that can be used to make good policy decisions regarding EE investments.

Issue 5: Baseline Issues in NTG

The determination of the baseline against which to measure energy savings is one of the most important aspects of an EE evaluation. One of the stronger aspects of the DNV report is its discussion of baselines in Appendix B and specifically which baseline is appropriate for various situations. This is particularly important in the verification of gross savings, which was conducted as part of the custom program savings verification (CPSV) analyses. It is also important for self-report survey design and implementation. In the efficiency attribution parameter NTG assignment (discussed in Section 1.2), the survey tries to get at partial free ridership by asking the following question:

> *Without <the program>, would you have installed <measure> that was "standard efficiency on the market at that time," or "between standard efficiency and the efficiency that*

---

[17] New York State DPS/Office of Clean Energy, *Evaluation, Measurement & Verification Guidance*, November 2016, p. 10. http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/255ea3546df802b585257e38005460f9/$FILE/CE-05-EMV%20Guidance%20Final%20%2011-1-2016.pdf

[18] A version of this approach was taken in Violette, D.; Ozog M.; Cooney, K. (2003), Retrospective Assessment of the Northwest Energy Efficiency Alliance -- Findings and Report. Prepared for: Northwest Energy Efficiency Alliance, Ad Hoc Retrospective Committee, December 8. See: http://www.theboc.info/pdf/EvalBOC_SummittBlue_NEEA_2003.pdf

you installed?

The participant can only answer this question if they know what the standard efficiency is for alternative equipment to compare to the higher efficiency equipment installed by the program, and if they need to know what the efficiency levels might be if they make the investment several years into the future (i.e., responded that they would have made the investment at a later date). There is sometimes a concern that customers are responding based on their view that any new equipment simply exceeds the efficiency of the existing equipment. In this case, they may be just moving from the efficiency of their existing equipment to standard efficiency and are not exceeding standard efficiency levels. This can be a complex question, as industry standard efficiencies may be higher than the efficiency of the lowest cost replacement equipment. Respondents may be overly confident about the likelihood that they would have installed higher than standard efficiency equipment in the absence of the program. This may result in an inaccurate estimate of partial free riders.

Making sure that the respondents are able to accurately project what investments they might have undertaken in the absence of the program often requires the use of setup questions that help respondents recall the sequence of past events, how these events affected their decision to adopt the measure, and awareness of equipment options that represent standard efficiency as opposed to the lowest cost replacement equipment.

Additionally, as noted in Issue 1 above, there are baseline considerations in the calculation of early replacement and whether that is approached through gross or net savings. Again, the only way to really address concerns about potential biases in the response to NTG questions that have a baseline assumed (e.g., installing equipment above standard practice) is to perform sensitivity analyses.

Issue 6: Tailoring Attribution Studies

It is important to tailor attribution studies to the specific programs being evaluated. Programs may have unique elements in terms of program design and delivery, which can affect net savings. Addressing these program design and implementation features may require tailored evaluation approaches.

Self-report survey methods have been most often used for incentive-driven C&I programs. However, there are a growing number of programs that are trying to move away from paying out large incentives. Instead, there is greater emphasis on engagement, information, and business case development—all of which support a more favorable environment for investments in EE. For these programs, it is often important that the survey introduces the ways support was provided through the program. This would include making sure that program training, analysis, and support are described to the participant. These can be particularly difficult for the respondent to recall if the survey takes place 1 year or more after participation. A program driven by financial incentives to induce participation has one major event (i.e., the payment of incentives) that the respondent can easily recall. More sophisticated programs that work to engage and support customers in making EE investments can require different survey designs to capture these non-incentive influence factors.

It can be important to work with program implementation managers and experts in the market (e.g., utility account managers and trade allies) to develop hypotheses that can be explored through the survey effort. These would include hypotheses regarding the different influence pathways used by a program to reach customers—particularly if there is a goal to move away from or better support customers and reduce the need for high cost incentives and rebates.

Finally, one of the best practices identified in the Jurisdictional Review report involved gaining insight into the project story by working through events that led to the installation of equipment. The goal is to have participants talk about their project-related decision-making and what factors went into that process. This may be conducted for a smaller set of customers than the overall NTG sample, but these insights can be important when making judgments about attribution.

Review and Analysis of Net-to-Gross Assessment Issues for Natural Gas Demand Side Management Custom C&I Programs

August 25, 2017



Prepared by

research into action inc

3934 NE Martin Luther King Jr. Blvd., Suite 300 | Portland, Oregon 97212
www.researchintoaction.com

Review and Analysis of Net-to-Gross Assessment Issues for Natural Gas Demand Side Management Custom C&I Programs

August 25, 2017

Funded By:

Enbridge Gas Distribution, Inc.

Prepared By:

Research Into Action, Inc.

Ryan Bliss
Mersiha McClaren
Nathaniel Albers
Jen Loomis
Elizabeth Focella
Jane Peters

research into action inc

www.researchintoaction.com

PO Box 12312
Portland, OR 97212

3934 NE Martin Luther King Jr. Blvd., Suite 300
Portland, OR 97212

Phone: 503.287.9136
Fax: 503.281.7375

Contact:
Jane S. Peters, President
Jane.Peters@researchintoaction.com

Table of Contents

Executive Summary

As part of its mission to regulate Ontario's natural gas sector, the Ontario Energy Board (OEB) has developed guidelines specifying adjustments the natural gas utilities should make to gross energy and demand savings to estimate how much savings actually resulted from the programs' activities (that is, *net* savings). These adjustments include reducing savings accomplished through the program that would have occurred without program involvement (free-ridership) and adding savings caused by the program but without program participation (spillover). OEB also has produced guidelines on the allocation of savings to parties other than the program that may have influenced the energy-saving activities.

This report presents a review and analysis of literature relating to the adjustments described above, particularly as they relate to the natural gas utilities' custom C&I programs. This review and analysis demonstrates that many potential problems exist with the way that net savings assessment has been conducted. Particularly problematic are self-report methods, which are very common for their low cost and ease of administration. Such methods, however, can easily result in over-estimation of free-ridership for multiple reasons: respondent self-selection bias; a tendency to provide a "socially desirable" response to questions about what would have occurred absent the program; the tendency to rationalize past decisions as arising from internal motives; difficulty envisioning hypothetical alternatives; lack of awareness of all the factors that may have influenced an action.

Apart from the above issues – which limit the ability of a program participant to provide an accurate description of what would have occurred absent the program's influence – there are multiple methodological challenges to assessing net savings. A lack of statistical precision can produce estimates that may change notably from year to year. Spillover often is not included in net savings research and, when it is, it may very likely be under-estimated. Although OEB guidelines indicate that spillover should be accounted for in estimating net savings, OEB currently does not approve a spillover adjustment to the natural gas utilities' gross savings from custom commercial and industrial (C&I) programs.

Several policy considerations relate to how net savings are defined and assessed. For one, applying variable and unpredictable net savings adjustments retroactively can lead to conflict and litigation from dissatisfied shareholders (Kushler, Nowak, and Witte 2014). While it might be bad policy to settle for a clearly inaccurate net savings assessment to avoid such conflicts, it would be reasonable policy to search for an approach that is defensible and avoids conflict.

Another policy issue is whether the value of conducting net savings research on a regular basis justifies the cost. Some evaluators (e.g., Violette et al. 2015) have concluded that it may not, even when the research is conducted with relatively inexpensive self-report methods.

At least partly as a result of one or more of the above issues, recent years have seen strong trends toward estimating net savings by applying a negotiated (also called "deemed" or "stipulated") net-to-gross (NTG) ratio to gross savings (Kushler, Nowak, and Witte 2012, 2014; SBW, Research Into Action, Inc., New Horizon Technologies, Inc., and Ridge & Associates. 2013).

A final policy issue this report relates to is OEB's requirement to establish a method to allocate some energy savings from program-funded projects to other parties that might have influenced those projects (Ontario Energy Board 2014, pp. 21-22). Our reading of the requirement is that such savings should be allocated out of the program's *net* savings – that is, after adjusting for free-ridership and spillover. We argue that such an approach is inconsistent with the definition of free-ridership, which should include the influence of such other parties. If the allocation is done after the application of free-ridership and spillover adjustments, then the utility is penalized twice for the effects of the same external influences.

Based on our review and analysis, we offer the following recommendations to OEB and the natural gas utilities:

- **Develop a negotiated (also called "deemed" and "stipulated") NTG value.** This value should be based on a range of inputs, including a review of researched NTG values from similar programs in comparable jurisdictions that account for free-ridership and spillover, at a minimum, but also market effects if possible. Assessment of applicable NTG values from multiple studies should not treat all inputs equally but should follow a meta-analytic approach, which includes reviewing the study quality, assessing study heterogeneity, and developing a pooled estimate of variability based on the variabilities reported in the studies. The pooled estimate is a better representation of what the true estimate is in the population and it can provide insight into variability around NTG that are important to consider when determining what the value should be. Part of reviewing study quality should include assessing efforts taken to reduce the self-report biases identified in section 3. Other inputs to the negotiated NTG value should include structured expert judgment and any available market data or macroeconomic analyses. In developing the negotiated value, it may be valuable to employ a "value of information" approach, such as described by Violette et al. (2015).

- **Allocate any savings to parties other than the program only from the free-ridership portion of gross savings.** By definition, free-ridership represents the program-claimed savings that would have occurred without program assistance, which must include savings attributable to other parties. Allocating savings net of free-ridership to other parties doubly penalizes the program.

As noted in the body of this report, establishing a negotiated NTG value does not preclude doing NTG research, as such research may be valuable for program planning and implementation as well as to inform periodic adjustments to the negotiated NTG value. We recommend that OEB and the natural gas utilities observe the following when NTG research is conducted:

- **Always include spillover and, if feasible, market effects assessments.** As documented in the body of this report, failure to account for these factors will underestimate NTG.

- **If using self-report, employ methods to reduce the bias toward high free-ridership.** Energy Trust of Oregon, with input from Research Into Action, Inc., developed an approach to free-ridership assessment that attempts to control for the high-free-ridership bias of other self-report methods in addition to reducing customer fatigue (see Bliss, McClaren, Folks, and Kociolek, 2015; Roy and Bliss 2012). This alternative approach balances the counterfactual assessment with a component that assesses the influence of the various program interventions, which typically produces a lower free-ridership estimate than the counterfactual (PWP and Evergreen Economics 2017).

- **Assess free-ridership as close as possible to project implementation.** The longer the time that has elapsed between the implementation of the project and the assessment of the decision-making that went into the project, the less salient the external influences (including the program influence) will be to the program participant and the more likely that participant will be affected by the biases toward free-ridership responses.

- **Use multiple methods and triangulate the NTG estimate.** The use of multiple methods, such as surveys of contractors as well as program participants, is now generally regarded as best practice among energy efficiency experts (Kushler et al. 2014; PWP and Evergreen Economics 2017).

Following the above recommendations may allow the natural gas utilities to continue offering large C&I customers in Ontario opportunities to generate high energy savings through custom programs that may not otherwise be achievable.

1. Introduction

The Ontario Energy Board (OEB) regulates Ontario's natural gas sector. As part of its mission, OEB has specified that the natural gas utilities should adjust gross energy and demand savings totals by free-ridership (energy savings accomplished through the program that would have occurred without program involvement) and spillover (energy savings caused by the program but occurring without program participation). OEB also has produced guidelines relating to the allocation of savings to parties other than the program that may have influenced the energy-saving activities. The purpose of these adjustments is to estimate programs' *net savings*, or the savings that actually resulted from the programs' activities.

Currently, OEB approves adjustments to the natural gas utilities' gross savings from custom commercial and industrial (C&I) programs based on researched free-ridership but not spillover.[1] This report presents a review and analysis of literature relating to net savings estimation to shed light on OEB's guidelines and requirements as they relate to the natural gas utilities' custom C&I programs. The report argues that estimating net savings through annual research is problematic for multiple reasons and argues instead for establishing a negotiated (also called deemed or stipulated) net savings approach for custom C&I programs.

1.1. Background

The Ontario Energy Board (OEB) regulates Ontario's natural gas sector. As part of its mission under the Framework for natural gas demand side management (DSM; see OEB 2014a), OEB developed filing guidelines for natural gas DSM programs (OEB 2014b). Among other things, these guidelines identify adjustment factors to be applied to the gross energy and demand savings totals reported by DSM programs to "ensure that the energy savings that are the result of DSM programs truly reflect those which the gas utilities directly influenced" (p. 20). Those adjustment factors include free-ridership and spillover. They also include attribution, which the OEB explains as relating "to whether the effects observed after the implementation of a natural gas utility's DSM activity can be attributed to that activity, or at least partly results from the activities of others" (p. 21).

In May 2015, the natural gas utilities contracted with DNV KEMA (now DNV GL) to carry out a study of free-ridership for their custom commercial and industrial (C&I) programs (Ontario Energy Board 2015a).[2] However, this research, now under OEB management, addresses free-ridership only, and not spillover (Ontario Energy Board 2015b). As of the preparation of this report, the results of the DNV GL evaluation have not been made public.

1 In fact, the Evaluation, Measurement, and Verification (EM&V) plan (DNV GL 2016) included a plan to conduct spillover research, but OEB determined there would not be sufficient time to complete the spillover research (Marc Hull-Jacquin, Enbridge Gas Distribution, personal communication). Note that the plan was to collect spillover data only through a participant survey. As argued in section 4.2.2 of this report, such an approach likely would underestimate spillover.

2 This study was undertaken with the endorsement of the Ontario Natural Gas Technical Evaluation Committee (TEC). In August of 2015, OEB announced a plan to transition the TEC's evaluation activities to OEB under the new DSM evaluation governance structure.

1.2. Purpose and Organization of This Report

The purpose of this report is to review and analyze literature on net savings estimation as it relates to OEB's guidelines and requirements and offer a recommendation on an appropriate method to be applied going forward. Section 2 of the report briefly summarizes the various approaches to assessing net savings (the savings that resulted from program activities) and developing a NTG ratio – the ratio of a *net savings* to *gross savings*, or the total savings that occurred through program-funded energy efficiency activities. The remainder of the report then presents information from a wide range of sources that brings into question whether OEB's selected approach accurately assesses the savings that resulted from the natural gas utilities' custom C&I programs' activities.

Section 3 focuses on the challenges inherent in the use of customer self-report survey data to assess free-ridership. This is the most common free-ridership assessment approach because of its low cost, and it is the approach that was used to estimate free-ridership and NTG for the Ontario natural gas utilities' custom C&I programs. Such challenges include several well-researched and -documented psychological tendencies as well as research design and implementation practices that would tend to exaggerate free-ridership values. This section argues that such challenges may particularly affect assessment of free-ridership in custom programs.

Section **Error! Reference source not found.** presents information on why – apart from the above challenges to the self-report methodology – researched NTG values likely are not accurate in any given year. Section 5 then discusses the policy issues related to the identified research limitations. These include the weighing of the cost of NTG research against the value of that research and the conflicts that may arise when researched NTG is retroactively applied to a program's gross savings. Section 5 also discusses how the logic behind NTG assessment relates to OEB's requirements regarding the attribution of energy and demand savings to parties other than the program.

Following the above sections, we present a brief conclusion and our recommendations to OEB and the natural gas utilities.

2. Review of Net Savings Assessment Methods

Evaluators are often required to calculate a program's net savings by applying net-to-gross (NTG) adjustments to the gross savings. Evaluators use a variety of methods to estimate NTG (Violette and Rathbun 2014), but our review of the literature reveals that the industry largely recognizes free-ridership and spillover to be the primary components of NTG estimation.[3]

Free-ridership (free-ridership, FR) refers to the portion of energy savings that participants would have achieved in the absence of the program through their own initiatives and expenditures (Violette and Rathbun 2014). Free-ridership ranges from 0 to 1, with 0 being no free-ridership (or, total program attribution) and 1 being total free-ridership (or, no program attribution). The values in between represent varying degrees of partial free-ridership. Spillover (SO) refers to the program-induced adoption of measures or actions by non-participants and participants who did not receive financial incentives or technical assistance from the program (Violette and Rathbun 2014). Spillover ranges from 0 to infinity, with 0 being no spillover and values greater than 0 demonstrating the existence and magnitude of spillover. Evaluation teams use the following formula to calculate a NTG ratio when relying solely on these components:

$$NTG = 1 - FR + SO$$

The following sections review some of the more common methods for estimating NTG.

2.1. Self-Report Surveys

Self-report survey is the most commonly used method for estimating NTG for those programs that target consumers directly and for which it is not possible to randomly assign consumers into a control and experimental groups. Our review of the literature reveals considerable variation in how evaluators and jurisdictions estimate NTG via self-report surveys – both in the questions asked and the algorithms used to estimate free-ridership and spillover. In the following sections, we report some basic tenants of the self-report survey method.

2.1.1. Free-ridership Estimation

Evaluators typically survey program participants to estimate free-ridership, but some evaluators conduct surveys with market actors (such as program-affiliated contractors) to inform free-ridership estimates (Violette and Rathbun 2014). To estimate free-ridership, evaluators typically ask survey respondents about what would have happened in absence of the program (the "counterfactual") and/or how much influence the program had on the upgrade decision (Violette and Rathbun 2014). Evaluators may ask

[3] Although some jurisdictions incorporate leakage and market effects when considering net impacts (Messenger et al. 2010), these components are rarely estimated. Market effects are changes in the adoption of energy-efficient products, services, or behaviors due to program or policy interventions. Leakage refers to indirect or unintended program effects. For example, if a program provides a discount for an LED at retail stores to increase LED adoption in the residential sector, some of those discounted bulbs could "leak" (be installed) in the nonresidential buildings because contractors are buying them.

participants to assess the counterfactual or program influence regarding their upgrade project as a whole or may ask participants about each specific measure or groupings of measures. Specific question and scoring design varies considerably in the industry. However, the industry is unanimous in the theoretical minimum of 0, or 0% free-ridership, and a maximum of 1, or 100% free-ridership.

The measure- or program-level free-ridership value typically is calculated as the mean of the sample values from the self-report research, often weighted by the total savings of the sampled projects.

2.1.2. Spillover

Evaluators often use self-report surveys to estimate both participant and non-participant spillover. Participant spillover refers to program-attributed savings from additional non-incented measures installed by participants who were influenced to do so by their experience participating in the program. Non-participant spillover refers to program-attributed savings from measures installed by non-participants who were influenced to do so by either directly or indirectly by the program.

Evaluators may survey program participants and non-participants to estimate spillover or may survey market actors (such as program-affiliated contractors) to inform spillover estimates. Evaluators use a variety of survey techniques to gather information on the measures installed outside of the program and the relative program influence on said measures. Evaluators may use primary or secondary research to estimate savings values for measures installed outside of the program.

Not all energy savings from measures installed without program incentives count as spillover. A common approach is to determine the amount of savings to attribute to the program based on the level of program influence on the decision to install the measures, as assessed from the surveys with participants, non-participants, or market actors. One approach is to establish a threshold level of influence and count all the savings from an installed measure if the rated program influence exceeds that threshold. Another is to attribute a portion of the savings for a given measure based on the rated influence. For example, a rated program influence of "3" on a 1-to-5 influence scale (from "no influence" to "great influence," say) might result in attribution of 50% of the savings to the program, while a rated program influence of "1" might result in 0% attribution and "5" might result in 100% attribution.

While self-report approaches to free-ridership yield a free-ridership percentage for each respondent, self-report spillover research typically yields a total spillover energy (or demand) savings value for each respondent. The measure-, project-, or program-level spillover percentage is calculated as the total spillover savings divided by the total measure, project, or program savings.

2.2. Experimental Approaches and Billing Analyses

Randomized control trials (RCT) or quasi- experimental methods (QEM) rely on billing data for estimating net savings. The distinction between the two is that RCT allows random assignment of customers to treatment and control groups while QEM may use a control group that is not randomly selected or, in some cases, does not even use a control group. Both methods typically use before-and-after-program billing data from the treatment and control groups to assess program effects, often attempting to control for other factors, such as weather. Both methods generally require large samples and selection of an appropriate control group and can be costly to carry out. Incomplete billing data can contribute to the challenge of conducting this type of analysis.

An RCT approach, which is recommended by the State and Local Energy Efficiency Action Network for behavior based programs (SEE Action 2012), will produce an estimate of net energy savings that is internally valid and unbiased, but it is not always feasible to implement (e.g., one cannot randomly assign subjects to naturally occurring groups). Naturally occurring groups occur when the program is an open-enrollment or opt-in program. Most custom C&I programs are opt-in programs.

The key challenge of opt-in program is self-selection bias. Self-selection bias refers to pre-existing differences (e.g., building square footage) between those in the experimental and control groups. The selection bias can be minimized through the use of statistical methods for sampling such as "regression discontinuity"[4] or "matched controlled group"[5] (SEE Action 2012 and Hall et al. 2004). However, the heterogeneity of large C&I custom participants makes matching a challenge. Moreover, unless Advanced Metering Infrastructure (AMI) is in place, billing data are not likely to be sufficiently granular to see the effect.

2.3. Market Sales Data Analysis

Sales data analysis is another method for estimating free-ridership and various components of spillover. The most common approach involves cross-sectional comparisons of sales of energy-efficient products or services in the area served and not served by the program. For example, efficient water heater sales in Ontario could be compared with efficient water heater sales in other areas of Canada, including regions with and without water heating programs. Water heater purchases in a specific time period serves as the dependent variable in a regression-based model. Independent variables in the model can include elements of program support, water heater technology saturation at the beginning of the time period, the length of prior program support in the area, and household-level measures of demographic, economic, or social characteristics.

The primary challenge is the selection of an appropriate comparison area and the availability of market sales data. The regression does reduce the need for a perfect comparison area as demographic and social characteristics can be adjusted for. Nevertheless, this method suffers from omitted variable bias – that is, the regression will likely not be able to account for all influencing factors.

2.4. Top-Down or Macroeconomic Modelling

Evaluators can rely on top down or macroeconomic models of sector-level state, regional, or national data on programs and target markets to estimate net impacts. Such models are based on changes in aggregate energy consumption (rather than changes in consumption for a specific account, as analyzed in billing analyses) as a function of energy efficiency efforts. Such analyses require a standardized measure of energy efficiency "effort" (e.g., program expenditures) as well as sophisticated modeling to identify the impact of a given program year's efforts over several succeeding years.

4 The regression discontinuity method selects a group of households just below the energy usage cutoff level as the control group and a group of households just above the energy usage cutoff level as the treatment group.

5 The matched control group method selects a control group with demographic and usage characteristics similar to those of the treatment group. The Regional Technical Forum (2010) recommends that, at a minimum, home type, location, and total baseline consumption characteristics of the control group should be similar to those of the treatment or experimental group.

2.5. Structured Expert Judgment

Some jurisdictions rely on a panel of experts to provide information used to calculate NTG. In these jurisdictions, a panel of experts knowledgeable about specific technologies and markets are asked to estimate baseline market share or to forecast market share, assuming common facts about the program, technologies, and other factors. In the Pacific Northwest, the Regional Technical Forum (RTF) helps utilities determine deemed savings values that take into the account baseline conditions, which includes free-ridership and spillover. The RTF uses an advisory committee, composed of regional experts, and subcontractors to regularly develop, update, and review a list of energy efficiency measures and determine appropriate deemed savings values based on engineering and market research.

2.6. Negotiated or Deemed Values

Deemed, stipulated, or negotiated values are NTG ratios that the program or commission determines are applicable and reasonable to apply to a program or portfolio. The NTG value deemed acceptable by the commission may come from a variety of sources, including:

- Literature review of other NTG studies from similar jurisdictions
- Structured expert judgement
- Market sales data analyses
- Top-down or macroeconomic models of data on programs and target markets
- Engineering estimates

Typically deemed values are adopted for consumer-facing or downstream programs. They typically are employed to save money and time compared to conducting monthly or annual research to determine NTG values, but they may also be used to avoid arguments concerning the calculation and award of utility shareholder incentives that may occur when researched NTG estimates are applied retroactively to gross savings estimates (Kushler, Nowak, and Witte 2014). We discuss these motives for using negotiated values in more detail in section 5.2.

Many jurisdictions rely, at least partially, on deemed values. To arrive at the deemed values, jurisdictions may use evaluations of programs and measures that include assessments of free-ridership and spillover. These evaluations may use some combination of the aforementioned methods to determine NTG and then, rather than conducting NTG research monthly or annually, rely on the deemed NTG values for a longer period of time. The jurisdictions revisit the deemed NTG values on some predetermined research schedule or when some element of the program changes or the market appears to be shifting somehow. To save money and resources, about 70% of all states apply deemed values determined from other jurisdictions' research (Kushler, Nowak, and Witte 2012).

One potential tradeoff of using the deemed approach is the lack of insight deemed values give program planners about how the market may be changing over time. In many cases, jurisdictions will allow the application of the researched NTG values for some programs or measures and apply deemed values to other programs where they are less concerned about insights into the market. As discussed in greater detail in section 5.3, some jurisdictions use deemed NTG values (or base compliance on gross savings, which is logically equivalent to having a deemed NTG value of 1.0) but also require NTG research to inform program planning.

3. Disadvantages of Self-Report Method

As noted in the previous section, self-report is the most common method for NTG estimation for downstream incentive programs, including C&I custom programs. It is the primary approach used to assess NTG among such programs in Ontario. The limitations of self-report to assess free-ridership, and consequently, NTG, are numerous, and several have received considerable attention in the literature (Peters and McCrae 2008; Ridge et al. 2009). It is important to note that the limitations of self-report are problematic not just for estimating free-ridership but for survey research more generally. The limitations may be organized into three broad categories: factors limiting the ability to respond accurately, research design and implementation issues, and factors specific to custom programs that would tend to exacerbate the effect of the other limitations.

3.1. Factors Limiting the Accuracy of Responses

Psychological research provides numerous reasons for why the responses people provide on self-report measures should be interpreted with caution. Below, we describe several pertinent and well-researched theories that highlight the potential for inaccurate self-reporting. While these tendencies and biases are found to reduce the accuracy of responses, it is important to note that they do not suggest that respondents are entirely unable to notice the program's influence, nor do they mean that there are no respondents who are not completely clear-eyed about their own motives and external influences. Rather, the research suggests a tendency to obscure respondents' judgment in the aggregate, with a resulting impact on the evaluation of program attribution.

3.1.1. Difficulty Estimating and Reporting Attribution

To assess free-ridership, respondents may be asked whether they, or the organization they represent, would have engaged in the energy efficiency behavior had the program not been in place. They may also be asked to account for what specifically caused them to decide on this energy efficiency action. In other words, respondents are tasked with determining the correct attribution for their behavior – who gets credit for the actions they took. Decades of research have documented that the attributions we make for our and others' behavior are often incorrect or at the very least, do not recognize the range of factors that lead to a given behavior.

Research suggests that a variety of motivations – the desire to maintain consistency between attitudes and behavior, to see oneself in a positive light, or to present oneself in a positive light to others – might all contribute to inaccurate or limited accounts for behavior (Kunda 1987). This means that when respondents (those that have opted for the energy efficiency behavior) are asked about the reasons for their behavior, their motivations will likely bias how they respond.

For example, the motivation to maintain consistency between attitudes and behavior suggests that respondents might infer that since they engaged in the energy efficiency behavior, they must in fact have favorable attitudes toward energy efficiency. This would bias them to reason that, since they have positive attitudes toward energy efficiency, they would likely have engaged in this behavior regardless of the program. This would overestimate free-ridership.

Relatedly, people tend to take credit for their successes and explain away their failures (Miller and Ross 1975). This self-serving bias suggests that if the energy efficiency behavior elicited positive outcomes for the respondent, they would be even more likely to believe that the success rests on their decision as opposed to something external to themselves. Thus, they would attribute the decision to themselves and subsequently believe that they would have engaged in the behavior even if the program did not exist. Ultimately, this would overestimate free-ridership. Additionally, when interviewed by an evaluator of an energy efficiency program, respondents might be nudged to attribute their behavior to their, socially desirable energy efficiency-positive attitudes, a tendency which we describe next.

3.1.2. Difficulty Reporting the Hypothetical Alternative (Counterfactual)

When respondents are asked whether they would have engaged in the energy efficiency behavior without the program, they are being tasked to imagine an alternative reality. Without having been in that situation, they are asked to imagine what they would have done if the program in question, that was designed to promote energy efficiency, never existed. This is asking the respondent to imagine the hypothetical with the hope that their speculation leads to an accurate assessment of their assumed behavior. Not only do they need to imagine a fictitious scenario, they then must imagine what their behavior would have been. To construct this alternative reality, respondents need to speculate, drawing from any information that may be available to them. This act of imagining would be influenced by numerous factors including what is salient to them at the time of the interview (energy efficiency is likely at the top of their mind), as well as the biases (attribution bias, the tendency to rationalize past decisions) we discussed in this section -- all of which should lead the participant to say they would have done the energy efficiency behavior regardless of the program and, consequently, lead to an overestimate of free-ridership.

While solutions are provided including by Ridge et al. (2009) and Violette and Rathbun (2014), the proposed solutions may simply increase the chances of arbitrariness in the free-rider score calculation, a topic which we will discuss in more depth later in this section.

3.1.3. Tendency to Rationalize Past Decisions

Because people prefer consistency, when they are made aware that their actions do not align with their attitudes, they experience a basic feeling of discomfort known as cognitive dissonance (Festinger 1957; see also Stone et al. 1997). Notably, this desire for people to make their behavior consistent with their attitudes has been applied to encouraging environmental sustainability (Dickerson et al. 1992). Pertinent to our discussion, when a person is asked to imagine whether they would have engaged in the energy efficiency behavior had the program not existed, they may be faced with a conundrum. Given that they have already, publicly, done the energy efficiency action, if they express an attitude inconsistent with their behavior, their attitudes would be out of step with their behavior, and subsequently cause them discomfort. The easiest route to reduce the dissonance should be to bring one's attitudes in line with their energy efficiency behavior. Thus, this would cause the person to change their attitudes to be more positive to energy efficiency, which would make it more difficult to imagine a world in which they would not have engaged in that behavior to begin with. Essentially, the avoidance, or attempt to resolve, dissonance should bias the respondent to say they would have engaged in the behavior regardless of the program (Peters and McCrae 2008).

3.1.4. Tendency to Provide Social Desirable Responses

Another potential limitation to self-report methods is the tendency for respondents to provide answers that are socially desirable (termed the "social desirability bias"). For example, to assess free-ridership, a respondent who indicates they have performed the energy efficiency behavior would then be asked (through either a single question or a series of pointed questions) whether they would have engaged in the energy efficiency behavior if the program did not exist. A responded who says "yes, I would have done the same energy efficiency behavior without the program" would be considered a free-rider. Psychological literature presents multiple reasons for why the response should be interpreted with caution, at the least. For one, the response to the question could simply be due to the possibility that the respondent wants to provide the socially appropriate answer, which would be that the energy efficiency behavior is the "right" thing to do, thus, it would be adopted by the respondent even if the program never existed.

Researchers and evaluators have proposed several solutions to address the likely possibility that respondents will be biased toward providing the socially desirable, though potentially untrue, response (see Ridge et al. 2009; Keating 2009). One of these solutions is to use a questionnaire where the "right" or socially appropriate answer might not be so obvious to the respondent; the California method seeks to do this. Another way to mitigate the social desirability bias is to ask multiple questions that may converge on a true estimate of free-ridership. Ridge et al. (2009) identified research on various for minimizing bias, which they believe will mitigate potential problems. They further noted a potential countervailing bias to exaggerate the influence of the program to help ensure that the program incentives continue.

While incorporating the various techniques that Ridge et al. (2009) mentioned may help, doing so lengthens the questionnaire, which adds other concerns, including increasing respondent fatigue (and potentially loss of engagement), and increasing cost of administering the survey. It also may make calculating a final free-ridership estimate more arbitrary, which we discuss in more detail at the end of this section.

Further, while these solutions are elegantly defended and may mitigate some of the contribution of the social desirability bias on the estimate of free-ridership, even accounting for this phenomenon does not remove the impact of other psychological phenomena and biases on self-report. These other biases also suggest the limitations of self-report and argue for caution when using this methodology, especially to assess the presumed impact of intentions on behavior. Below, we describe each of these documented biases and psychological phenomena and how they obscure an accurate estimation of free-ridership.

3.1.5. Failure to Recognize All Direct and Indirect Pathways of Program Influence

It is conceivable that the individual respondent may be unaware of all direct and indirect pathways of program influence. Primarily, when accounting for their energy efficiency behavior and assessing whether they would have engaged in the energy efficiency behavior without the program incentives, they may fail to recognize all the pathways of program influence and erroneously conclude they would have engaged in the behavior even if the program has not existed. For example, while respondents may note the influence of contractors or equipment vendors (who may be salient to respondents since they may have interacted to set up the energy efficiency solution), they may not recognize the degree to which the program influenced those trade allies. Thus, they may not fully appreciate the degree to which the program indirectly influenced their adoption of the energy efficiency behavior (Bliss, Sage, and Diebel 2017). Respondent tendencies to neglect these indirect pathways of program influence on their decision to opt for energy efficiency would thus inflate the free-rider estimate.

3.1.6. Difficulty Isolating Program Influence from Longer-Term Market Transformation Effects

The decisions and behaviors of people and organizations are not solely influenced by an individual program, but by a variety of other forces. As described by Vine et al. (2010), numerous public policies and market interventions influencing energy efficiency often operate simultaneously, and it is likely impossible to extract the influence of a single program. This is an especially difficult task for a single respondent. For example, in addition to the specific program in question, public policy (e.g., state government messaging advocating for energy efficiency, tax credits for energy efficiency measures) as well as market interventions (e.g., media coverage of energy efficiency issues, other private-sector advertising) and other forces such as energy efficiency education in universities and other schools likely all exert their influence on the consumer's behavior. The individual respondent would conceivably have difficulty identifying the unique contribution of the program on their behavior apart from the other numerous influences, including market transformation effects.

3.2. Research Design and Implementation Issues

There are several issues relating to how surveys are designed and implemented that can affect accurate attributions of behavior, by exacerbating the psychological forces described above or by other means. Three such issues are response bias, survey timing, and arbitrariness in scoring free-ridership.

3.2.1. Survey Design and Response Bias

Good data are predicated on good survey design. The hurdles at this initial stage of research include response bias, more general issues related to sampling, and questionnaire construction. Most NTG research attempts to incorporate good instrument-design practices, such as avoiding double-barreled questions[6], making questions as clear as possible to respondents, and avoiding leading questions (e.g., "How satisfied are you with the program's generous incentives?"). NTG surveys may not be as likely to incorporate multiple-item scales, as advocated by Baumgartner (2013). Experienced NTG researchers also generally understand the importance of attempting to reach and interview a contact who (theoretically) can report knowledgeably on the decision to do the energy efficiency project in question.

One looming issue within the area of survey design, however, is response bias. Pertinent to our discussion, response bias may inflate free-ridership estimates. For example, in the case of a person or organization that participates in a program to encourage taking an energy efficiency action, to assess free-ridership we would want to know whether that organization or person would have taken the action

6 Double-barreled questions that do not allow the respondent to differentiate separate things in the response. For example, asking the respondent to rate satisfaction "with the program and its incentive" does not allow the respondent to indicate satisfaction separately for the program and for the incentive.

if the program had not existed. It is possible that those adopters who have less positive attitudes toward energy efficiency might not have engaged in the energy efficiency behavior without program incentives (i.e., not a free-rider) but may also be less likely to want to take a survey about this behavior. Conversely, those with positive attitudes toward energy efficiency behavior may have indeed engaged in the energy efficiency behavior regardless of the program (making them a free-rider) but should also be more likely to take a survey about energy efficiency behavior. Thus, this response bias would overestimate the number of free-riders.

3.2.2. Timing of Surveys

Several researchers (Schwarz 2007; Keating 2009; Peters et al. 2010; Violette and Agapay-Read 2016) have noted that the timing of surveys is particularly important to ensure the most valid responses. The longer the time that has elapsed between the behavior and the self-report about the behavior, the more likely the respondent is to forget their intentions, the motivations, and other influences on their behavior (even if the respondent had been aware of them at the time of action). Returning to attribution theory, the respondent's difficulty in accurately attributing their energy efficiency behavior is increased the longer the time between the energy efficiency action and the survey because the less obvious influencers on the respondent's decision and action fade in their memory. Further, with a longer amount of time between the behavior and the self-report, the more likely the respondent is to be influenced by other psychological biases. For example, research on the mere-ownership effect (Beggan 1992) suggests that people value an object more once they own it. Once an object is theirs (as a gift or after purchase), people are more favorable than when it was not their possession. Hence, one may imagine that the respondent has begun to value the energy efficiency product simply by possessing it. When asked if they would have done the energy efficiency behavior without the program, their ownership of the product should bias their ability to imagine themselves without it, and to increase the value of the energy efficiency product. The more that time has passed, the more difficult it may be to imagine oneself without that now-valued object.

3.2.3. Potential Arbitrariness in Free-Ridership Scoring Methods

Finally, some (Violette and Rathbun 2014) have noted that there is considerable arbitrariness in scoring methods to create free-rider estimates. By using a lengthy survey, combining open-ended and close-ended questions, and interview methods that point out respondent's inconsistent answers, the interpretation of the data from these questionnaires becomes largely dependent on the interpretation of the evaluator. Granted, if evaluations are using the same calculation, they should reach the same estimation of free-ridership, making their estimations reliable. However, their relative agreement does not necessarily indicate accuracy. Their estimation, though agreed upon, may still be incorrect, and therefore invalid.

3.3. Challenges Particular to Custom Programs

Haeri and Khawaja (2012) argued that no traditional approach adequately accounts for either free-ridership or spillover, especially for commercial, industrial, and new construction programs. Particularly relevant to the discussion here, they argued that self-report is especially problematic for assessing free-ridership in C&I programs because of the complex decision making involved in those types of projects.

If Haeri and Khawaja (2012) are correct, the issues they cite may be particularly a concern for custom programs. Moreover, custom projects often are larger and have a longer planning horizon than prescriptive projects. The longer planning horizons complicate assessments of the degree to which programs influence project planning, which could exacerbate the psychological forces that would tend to exaggerate free-ridership.

This added complexity would naturally muddy respondents' ability to accurately identify attributions for their energy efficiency behavior. With multiple forces influencing their behavior, and at different times, it would be especially difficult for the individual respondent, who has a limited perspective, to determine whether the program had its intended effect or to identify the factors that truly did influence their behavior. Particularly, as previously discussed, respondents have difficulty isolating program influence from market effects as well as differentiating all the direct and indirect pathways of program influence. Adding increased complexity to this already less-than-transparent situation may especially impede respondent's ability to answer accurately. Importantly, obscuring respondents' ability to answer accurately would likely nudge them to rely on their biases; they are unsure and need to rely on something to guide their judgements.

A concrete example may help illustrate the above point. Large C&I programs often work with larger customers over a long period of time – sometimes, for a decade or more – to identify and catalog available energy efficiency projects. In such scenarios, it is possible that, when a particular project becomes prioritized for implementation, the customer's staff retain knowledge of the project as an option but have forgotten that it was the program staff who identified it in the first place.

As decades of decision making as well as social psychological literature document, complexity and ambiguity increases the likelihood that people will rely on their biases to make judgments (Frisch and Baron 1988). In this case, their biases (e.g., social desirability bias, self-serving bias) will nudge them to say they would have taken the energy efficiency route regardless of the program and therefore, lead to an overestimation of free-ridership.

3.4. Summary

The above discussion provides several reasons why self-report surveys probably do not provide accurate estimates of free-ridership. Much well-researched and validated psychological theory indicates that self-report research may overestimate free-ridership, and the complexities of decision-making in custom C&I projects may make self-report a particularly problematic way to assess free-ridership for such programs. Our review of the literature, in both psychological theory and that specific to energy efficiency Evaluation Measurement and Verification (EM&V), found little argument and no evidence that self-report assessment under-estimated free-ridership. The one possible exception, as noted above, is that surveyed participants may explicitly exaggerate the importance of the program to help ensure the continued availability of the incentives. While this possibility cannot be dismissed out of hand, it must be weighed against all the well-documented psychological tendencies that would bias self-report in the other direction. While attempts at varying levels of success have been implemented to mitigate the issues and biases that may influence free-ridership estimates, the theory and research cited above suggests that they are likely leading to bias in one direction. That is, psychological biases and issues related to survey design largely lead to over (and rarely under) estimation of free-ridership.

4. Other Methodological Concerns with Researched NTG

The previous section provided several reasons why self-report research may overestimate free-ridership and, hence, underestimate NTG. Those are important considerations, but they are not the only arguments for using a negotiated NTG value. The following subsections document two key issues with relying on researched NTG:

- Lack of statistical precision can mean that the researched NTG in a given year may not be accurate.
- Spillover is a potentially important part of NTG, but it often is not sufficiently accounted for in researched NTG.

4.1. Researched NTG Can Lack Precision

The components of NTG – free-ridership and spillover – can vary greatly from year to year and across programs. While both the use of different assessment methodologies and differences in program implementation can contribute to differences in estimated NTG, a lack of precision in the individual assessments also contributes to the differences. The issue of lack of precision is important and worth a brief discussion before we proceed to the reported NTG findings.

4.1.1. The Meaning of Statistical "Precision" and "Confidence"

In statistics, "precision," strictly speaking, refers to the range of values that repeated samples from a given population will produce. Every sample produces an estimate of some characteristic of the population it is drawn from but, obviously, no two samples will produce the same exact estimate of that characteristic. Thus, calculating free-ridership in two samples of custom projects from the same program in the same program year will produce two different estimates of free-ridership for the program. A sample has high precision if most repeated samples of the same size, and drawn using the same methods, would produce estimates within a small range of values.

But what do we mean by "most" repeated samples or a "small" range of values? The meaning of "small" refers to the stated precision level and the meaning of "most" relates to the desired level of "confidence." When evaluators talk about precision, it is always in the context of the confidence level. In evaluation, we often seek 10% precision at 90% confidence at the program level. That is, we want a sample such that, if we continued to draw additional independent samples, 90% of those samples would produce an estimate that is no more than 10% higher or lower than the estimate our sample produced. (This often is interpreted as meaning that such a sample gives us 90% confidence that the true population mean is within 10%, higher or lower, of the sample mean. While many statisticians believe this is not strictly speaking true, it is a useful way to think of the results.)

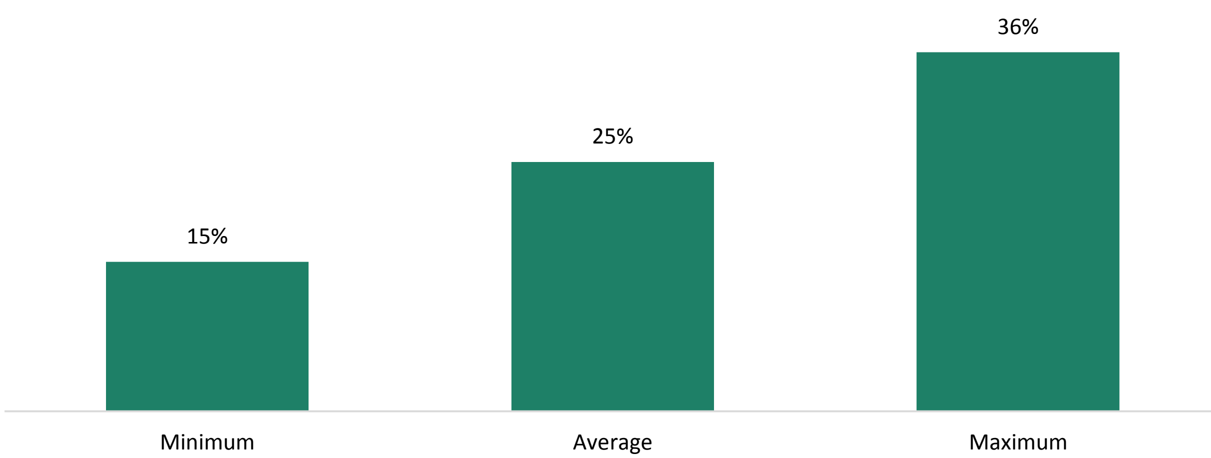
It should be clear, then, that even when samples are designed to produce 10% precision at 90% confidence, it is possible for two samples to produce noticeably different estimates of the same population value. One further point is important here, which is that a sample may be *designed* to produce 10% precision and 90% confidence but may not actually do so. This is because the level of precision is in large part a function of how variable the sample is with respect to the thing being measured – in this case, free-ridership. If most projects in the sample have similar levels of free-ridership, then there is low variability and good precision; but if the level of free-ridership is highly variable, then precision is not as good. Since the actual level of variability cannot be known in advance, researchers must base the sample design on the assumed variability. If that assumption is incorrect, then the assumed levels of precision and confidence also are incorrect.

4.1.2. Evidence of Variability in Researched Free-Ridership

The above background should help to put the following research findings in context. A review of free-ridership estimates across nine program types across multiple jurisdictions in the northwestern United States revealed notable variation in estimates across programs, in particular with custom programs (Cadmus 2017). This review of 13 custom C&I programs – seven in California, four in Oregon, one in New York, and one in Wisconsin – revealed a wide range of free-ridership estimates, from 11% for Energy Trust of Oregon industrial program in 2010 to 74% for a California Public Utility Commission (CPUC) agricultural custom program.

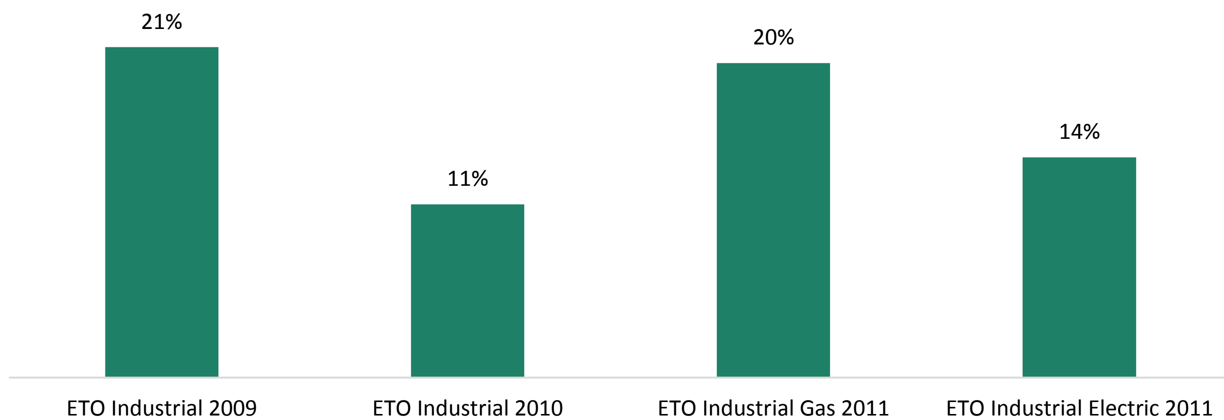
One program administrator, the CPUC, calculated the range of free-ridership for its agriculture and custom programs in 2009 to 2011. The values for the custom programs ranged from a low of 15% to a high of 36% (Figure 1). While this could reflect year-to-year differences in the programs' project make-up, it also likely reflects lack of precision in the estimates. Unfortunately, the report citing these values did not include estimates of precision, and the reference to the original source is no longer a live link. In any case, making policy or program planning decisions based estimates with so much year-to-year variability could easily lead to conflicting decisions.

Figure 1: Free-ridership Estimate Range for the CPUC's Agricultural and Custom Programs



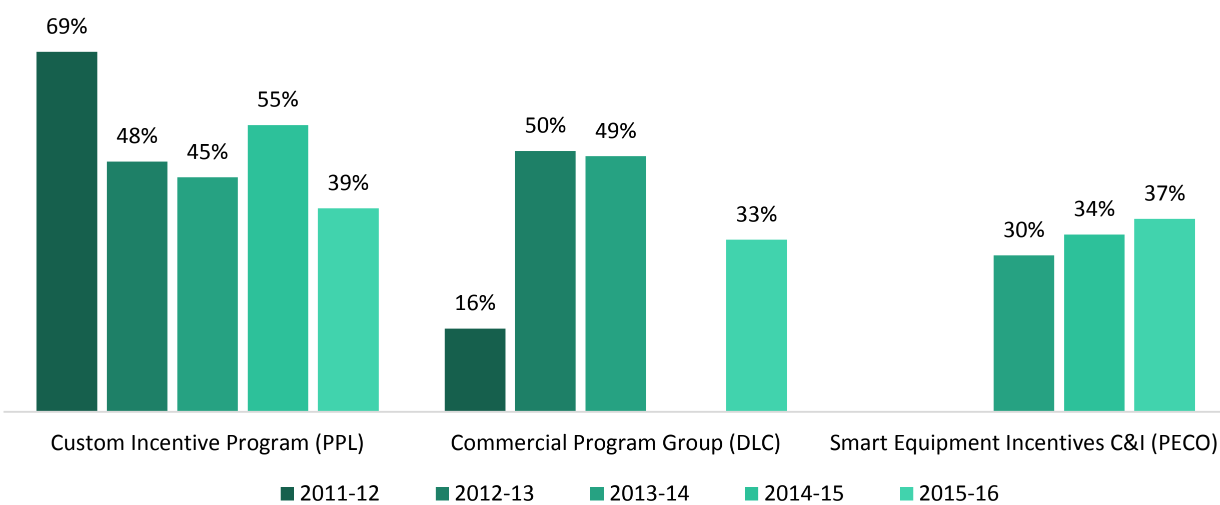
Examining a specific program's free-ridership values across multiple years sometimes shows variation that is difficult to interpret. For example, the free-ridership estimate for Energy Trust's Industrial program was 21% in 2009, dropped to 11% in 2010, and went back up in 2011 (Figure 2). Again, the year-to-year variability underscores the risk in relying on any specific estimate in determining the "true" free-ridership value.

Figure 2: Free-ridership Estimate for Energy Trust's Industrial Program, 2009-2011



Variability also existed in free-ridership estimates for C&I programs in several Pennsylvania utilities (Figure 3). As with the Energy Trust values, there was no clear pattern across utilities and years. For one program, free-ridership estimates trended down with a seemingly aberrant spike in the fourth year. For another, free-ridership tripled after the first year before falling to twice the starting point. For the third program, it slowly increased across years.

Figure 3: Free-Ridership Estimates for Pennsylvania Programs, 2011/12 to 2013/14



Sources: GDS Associates et at. 2014, 2015, 2016, 2017.

4.1.3. Exacerbation of Imprecision from Considering Spillover

The above subsection addressed free-ridership, but the lack of precision applies to spillover as well. In fact, spillover estimates may be even less precise than free-ridership estimates, as self-reported energy efficiency actions outside of efficiency programs are relatively low-frequency events. As Haeri and Khawaja (2012) point out, this means that a small *absolute* increase in spillover reported in a sample may result in a large increase in the spillover *percentage*.

Moreover, what is usually not considered is that a NTG estimate that includes both a free-ridership and spillover estimate (estimated separately) is not as precise as either the free-ridership or spillover estimate alone. That is because there are separate sources of variability for the free-ridership and spillover estimates that are combined when they are put together to form the NTG estimate.

Again, a slight digression into statistics is needed here. The precision of an estimate is a function of the *standard error* of that estimate. It is not necessary here to go into great detail about how the *standard error* is calculated, except to note that it is related to the *variance*, which is a measure of the variability of the sample component constituents – in this case, the individual free-ridership or spillover values that make up the sample – and to the sample size. When two estimates are combined, as when the separately estimated free-ridership and spillover are combined to estimate NTG, the variance around the combined estimate (the NTG in this case) is the sum of the variances of the components of that estimate (the free-ridership and spillover estimates).

Suppose, for example, an evaluation estimated free-ridership and spillover. Assume that samples of 68 observations generated estimated mean free-ridership and spillover values with 10% precision at 90% confidence. In both cases, the *variance* of the estimate is about .25, and so the *variance* of the NTG estimate is about .50, resulting in a precision of about 14% instead of 10%.

4.2. Spillover Is Not Sufficiently Accounted For

It is important to include estimates of spillover when free-ridership adjustments are made to ensure a balanced NTG ratio. Some evaluators have argued, and some regulators have accepted, that spillover and market effects balance out free riders (e.g., PWP and Evergreen Economics 2017; Khawaja, Haeri, and Hedman 2014; Haeri and Khawaja 2012). While there is as yet little empirical evidence for this argument, there is good theoretical reason to expect it is true. As an energy efficiency program succeeds in increasing trade allies' promotion of efficient equipment and end-users' recognition of the value of energy efficient investments, both self-reported free-ridership and spillover likely will increase (e.g., see Saxonis 2007). Yet, as documented below, not all states report spillover when estimating net savings. At the same time, current methods to estimate spillover may underestimate spillover savings.

4.2.1. Spillover Is Not Always Measured or Reported

When commissions/programs adjust gross savings by subtracting savings from free-riders, spillover should also be evaluated to provide for a balanced estimation of program effects (PWP and Evergreen Economics 2017; Kushler et al. 2014; and many others). One of the principles NEEP (2006) developed for estimating net savings is to "apply the concept of symmetry" which accounts for both positive (spillover) and negative (free-ridership) influences. Measuring free-ridership without accounting for spillover is not fully accounting for net program influences. Hence, retrospectively punishing programs for high free-ridership by reducing program-generated energy savings is considered “overly punitive” by some when the NTG ratio does not account for spillover (Khawaja, Haeri, and Hedman 2014, p.40).

Enbridge does not include spillover in its NTG analyses, which does not credit the program for the energy customers saved influenced by Enbridge programs (Synapse Energy Economics 2015). Participant and non-participant spillover is highly likely when programs have been in place for several years, as many prior participants are not still participating, yet were influenced to continue to pursue energy efficiency as a result (as in the case of Enbridge’s programs).

A growing number of states are recognizing the importance of including spillover estimates in their NTG ratios. Kushler et al. (2012) found that while 26 of the 39 states (67%) adjusted for free-riders, only 17 (44%) always included spillover. In a subsequent iteration of their survey, Kushler et al.(2014) found that 25 of 43 states include spillover (58%) and five more reported planning to. Table 1 displays which states adjust for free-ridership or spillover in their net savings, as reported by representatives in a phone survey.

Table 1: Reported Net Savings Adjustments by State*

Free-riders	Spillover	Number of States	States
Yes	Yes	33	Arkansas, California, Connecticut, District of Columbia, Florida, Georgia, Hawaii, Idaho, Kentucky, Maine, Maryland, Massachusetts, Michigan, Missouri, Montana, Nevada, New York, North Carolina, Oklahoma, Oregon, Rhode Island, South Dakota, Tennessee, Utah, Vermont, Wisconsin, Wyoming, and portions of Alabama, Georgia, Kentucky, Mississippi, North Carolina, and Virginia
Yes	No	4	Colorado, Illinois, Indiana, New Mexico
No	No	12	Arizona, Delaware, Iowa, Kansas, Minnesota, Nebraska, New Hampshire, New Jersey, Ohio, Pennsylvania, South Carolina, Texas
No data		2	

* Adapted from Kushler et al. (2014).

4.2.2. Spillover Is Easily Underestimated with Current Methods

In a report documenting the results of a recent literature review and expert interviews, PWP and Evergreen Economics (2017) note that estimated participant spillover usually falls below 5% of gross savings, while non-participant spillover estimates “vary widely.”[7] It may be more likely that self-report methods underestimate spillover. Underestimations of spillover can derive from a reliance on the survey respondents’ attribution of influence to the program. As discussed in Section 3, because of the tendency

[7] Although the authors report that some estimates of non-participant spillover exceed participant gross savings, such cases appear to be infrequent and may be limited to certain specific measure types, such as high-bay lighting (personal communication, Phil Degens, Evaluation Manager at Energy Trust of Oregon, August 15, 2017).

to rationalize past decisions, people tend to attribute energy efficiency decisions to themselves. This would be as likely lead to underestimation of spillover as overestimation of free-ridership.

Another reason is that self-report studies can identify only spillover activities done at the time of the survey. This is particularly problematic when a survey is conducted within the program year in which the respondent participated because it would not capture any spillover activities done after the survey but within the program year. To overcome this potential problem, a program might seek to conduct self-report surveys up to two years after program participation (Tetra Tech 2011). However, increasing the time that has elapsed between program participation and self-report surveys may lead to recall issues, with a resulting and greater tendency for the biases described in Section 3.

Bliss et al. (2017) argued that accurate survey-based spillover assessment must incorporate the perspectives of all parties involved in selling and installing energy efficient equipment – the equipment vendors, the installation contractors, and the end-users (program participants and non-participants). Specifically, in addition to assessing the program's direct influence on end-users, via marketing and outreach as well as learning the value of energy efficiency investments through program participation, accurate spillover assessment must assess the program's indirect influence on end-users via its influence on vendors and installation contractors. Accurate assessment of indirect influence must include assessment of: 1) the program's influence on the recommendations that equipment vendors and installation contractors make to their customers and on the recommendations that vendors make to contractors; 2) the equipment vendors' influence (through recommendations, stocking practices, and pricing) on installers; and 3) the vendors' and installers' influence (through recommendations, stocking practices, and pricing) on end users. Survey approaches that do not attempt to assess all those elements risk misestimating program influence.

Approaches that rely only on the end-users or the vendors and contractors, according to this view, cannot accurately estimate spillover because they cannot accurately assess both the direct and indirect pathways of program influence. While end-users are, at least nominally, able to identify program direct influence on their decisions (subject to the limitations identified in Section 3), they cannot report on the program's influence on vendors' and installers' practices, and so they cannot by themselves provide insights into program indirect influence. On the other hand, while vendors and installers can speak to the program's influence on their practices, they cannot report on the program's direct influence on end-users.

5. Policy Considerations and the Rationale for Negotiated NTG

This section discusses some of the policy considerations that proceed from, or are otherwise related to, the preceding discussion. First, some have concluded that the value of doing NTG research for a specific program year may not offset the cost. In addition, reliance on researched NTG, applied retroactively to gross savings, can generate conflict regarding the calculation and award of utility shareholder incentives. Following the discussion of the above issues, this section provides a summary of where negotiated or deemed NTG values have been used. Finally, this section addresses the related issue of how to attribute some portion of gas savings to parties other than the program in question and how that affects, if at all, the establishment of the NTG value.

5.1. The Value of Annual Primary Research May Not Justify the Cost

A primary reason for using NTG values is to accurately understand the amount of energy savings a program has generated so that policymakers can be sure ratepayer dollars are being spent in a cost-effective manner. However, conducting the studies uses a large portion of ratepayer dollars dedicated to EM&V (Messenger et al. 2010). Concerns over whether the funds spent on NTG self-report research justify the research costs, when deemed or negotiated values could be used instead, go back many years (e.g., Peters and McRae 2008; Messenger et al. 2010; SEEAction 2012) and continue to stimulate research and discussion (Violette et al. 2015; NEEP 2016).

Peters and McRae (2008) argued that funding self-report NTG research is not the most effective way of spending ratepayer dollars. Rather, research on motivations, behaviors, messaging, and intervention strategies may drive greater energy savings and would be a more cost-effective use of ratepayer funds. Two years later, though, this was still an open issue. In interviews with more than 80 energy efficiency experts, Messenger et al. (2010) found that those seeking more consistency in reporting impacts likely would encounter disagreement on using researched versus stipulated (deemed) NTG values.

SEEAction (2012) suggested that deemed NTG values are best used when "the expense of conducting NTG analyses and/or the uncertainty of the potential results are considered significant barriers." (p. 5-7) The authors caution that deemed NTG values are potentially less accurate than research-based approaches, but do not cite specific data to support that claim. A possible basis for that suggestion is the concern that deemed values should be based on comparison to "similar programs, *hopefully* applied to similar populations with a similar level of efficiency adoption and during a time period similar to that of the program being reviewed" (emphasis added). In other words, the potential for inaccuracy may come from basing the analysis on programs that do not have sufficiently similar populations, over a time period that is not sufficiently similar. Despite this note of caution, the authors suggest that conducting NTG research every few years and using those findings to stipulate NTG ratios for the intervening years is acceptable, "as long as the market influences and participants' behavior are relatively consistent" (SEEAction 2012, 5-7).

More recently, Violette and colleagues (Violette et al. 2015; NEEP 2016) have suggested that deemed or negotiated NTG values are sometimes close enough to the research-generated NTG value that policy decisions would be the same whether negotiated or original NTG values are used.

Violette et al. (2015) analyzed the costs and benefits of doing NTG research in Iowa compared to assuming a NTG value – in this case, a NTG value of 1.0. Specifically, the researchers compared the increased benefits of obtaining better NTG information to the cost of obtaining that information. The analytic model incorporated information on NTG values for similar programs in other jurisdictions to generate a distribution of probabilities for NTG values that differed from 1.0. The researchers then generated cost-benefit ratios under varying assumptions about research cost and rigor, research frequency, risk that true NTG departs from 1.0, and value of program design improvements resulting from NTG research. Under all scenarios, including ones with a low cost and high benefit of NTG research, the model indicated that the cost of annual NTG research outweighs the benefit for a custom C&I program, even compared to a deemed value of 1.0. Although the report does not consider the cost-benefits of NTG research compared to a deemed value of less than 1.0, it seems clear that it would weigh even more heavily in favor of the deemed value.

A guidance document on gross and net savings (NEEP 2016) expands on the earlier work by Violette et al. (2015). The authors of that document encourage utilities to consider the value of the information generated from NTG studies to determine whether the potential value/benefits of original NTG research outweigh the costs of conducting it. The authors recommend that policymakers consider the likelihood that original NTG research would produce information sufficiently different from current assumptions to result in program changes, and on that basis, consider whether updated gross savings and net savings information is needed to inform decision-making or whether spending ratepayer dollars on other types of research (e.g., market research) might be more valuable.

5.2. Reliance on Retroactive NTG Application Can Generate Conflict

Kushler et al. (2014) noted that conflict over net savings methods and results can arise – indeed, *has arisen* – when the results of net savings analyses have substantial financial impacts, such as on utility performance incentives or lost revenue recovery. Those authors noted:

> "Exacerbated by a policy of retroactively applying *ex post* estimates of free ridership, California degenerated into years of argument and litigation regarding the calculation and award of utility shareholder incentives." (p. 23)

Citing a study by the California Public Utility Commission (TecMarket 2010), Kushler et al. (2014) noted that the way in which NTG is calculated could mean the difference between nearly $400 million in earnings and a penalty of more than $100 million. Kushler et al. (2014) recognized that California's experience was an "extreme example," but even a less extreme experience can generate conflict. For those authors, avoidance of such conflict is one of the factors that has led to the "great proliferation" of deemed or negotiated NTG values in recent years.

5.3. Use of Negotiated NTG

Researchers observed strong trends among the US States in using deemed or negotiated NTG values for their programs or portfolios (Kushler et al.2012, 2014; SBW et al. 2013). In their review of 31 state's policies for estimating net savings, Kushler et al. (2012) found that 19 states use deemed values for their NTG ratios. Several of these states are listed in Table 2, all of which use deemed NTG values for their non-residential custom programs. Reportedly, regulators in Iowa and Arizona deem their NTG ratio at 1.0 because they "have accepted the argument that spillover and market effects balance out free riders" (Khawaja, Haeri, and Hedman 2014, 40).

Table 2: Use of Deemed/Stipulated NTG Values

State	Program(s) or Portfolio	NTG Value
Minnesota[a]	Portfolio	1.0
Arizona[b]	All programs	1.0
Iowa[b]	All programs	1.0
New Hampshire[c]	All programs	1.0
New Jersey[c]	All programs	1.0
New York[a,b]	All programs	.90
Michigan[a]	For all EE programs besides pilot, low-income, and education programs	.90
Hawaii[b]	All programs	.70

[a] Research Into Action, New Horizon Technologies, and Ridge & Associates (2013).
[b] Violette et al. (2015). Note that the information for Hawaii is not consistent with information in the Hawaii Energy 2014 Annual Report (Leidos 2014), which shows program-specific NTG factors and a composite NTG ratio of .78.
[c] Kushler et al. (2014).

Stipulated NTG values of 1.0 are common because many research studies estimating NTG factors have found that free-ridership and spillover roughly cancel each other out (Haeri and Khawaja 2012; Nowak and Witte 2014). Low-income programs and pilot programs targeting emerging technologies generally assume a NTG value of 1.0 because the target audiences demonstrate little free-ridership, as they are unlikely to purchase the newer, more expensive, energy-efficient products on their own.

In addition to the jurisdictions that explicitly identify a deemed NTG value, there are other jurisdictions that may require or encourage NTG research to inform program planning but do not apply NTG to assessments of program savings. In other words, these jurisdictions pay attention only to gross savings, not net, which is logically equivalent to having a stipulated NTG value of 1.0. For example, the Pennsylvania Utility Commission bases compliance with energy and demand reduction targets on gross verified savings, but it nevertheless requires Pennsylvania electric distribution companies to conduct NTG research to inform program design and implementation (GDS, Research Into Action, and Apex 2017). Similarly, as noted elsewhere in this report, utilities in the Pacific Northwest use deemed savings values that take into the account market baseline conditions, which includes free-ridership and spillover. In this case, gross reported savings based on the deemed values are net of free-ridership and spillover. Yet many of those utilities continue to conduct NTG research to inform program planning and implementation (e.g., Roy et al. 2016).

5.4. Attribution of Savings to Other Parties

Finally, it is important to clarify how the above relates to the discussion of "attribution" in the Ontario Energy Board's *Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2000)*. Section 7.2.2 of that document addresses "whether the effects observed after the implementation of a natural gas utility's DSM activity can be attributed to that activity, or at least partly results from the activities of others" (Ontario Energy Board 2014, 21).

The guidance presented in that section addresses two topics. The first – "attribution between rate-regulated natural gas utilities and rate-regulated electricity distributors" – is not relevant to this report, which is concerned only with the attribution of gas savings. Of concern to the present discussion is the second topic – "attribution between rate-regulated natural gas utilities and other parties (e.g., non-rate-regulated entities such as agencies and various levels of government, non-rate-regulated private companies, etc.)." Such other parties might include GreenOn, the Ministry of Environment and Climate Change, and any other large funding body that promotes energy efficiency in Ontario.

The *Filing Guidelines* state that natural gas utilities should establish partnership agreements with such other parties before program launch, specifying the shares (percentages) of natural gas savings to be allocated to the natural gas utilities and the other parties. If the percentage allocated to a given natural gas utility exceeds its percentage of total dollars spent by more than 20%, the utility should provide an explanation for the difference.[8]

Some attention has been paid to the question of sharing credit for energy savings among multiple influences (e.g., Skumatz and Vine 2010), although we identified no reports detailing a methodology for doing so. The important consideration for this issue, however, is whether the above reference to the allocation of natural gas savings refers to *gross* or *net* savings. The discussion in the introductory paragraphs of section 7.2 of the *Filing Guidelines* suggests that it refers to the *net* savings, as defined in section 2 of this report. Specifically, those sections refer to applying "attribution" as an adjustment factor separate from free-ridership and spillover. This seems to imply that the "attribution" adjustment, as defined above, would occur after adjusting for free-ridership and spillover.

We believe that such an approach is inconsistent with the meaning of *gross* and *net* savings as universally used in the energy efficiency evaluation community. In particular, it is inconsistent with the definition of free-ridership as the program-claimed savings that would have occurred without program assistance – meaning that net savings are those that occurred *only because of the program's assistance*. Another way of stating this is that the counterfactual in freeridership assessment theoretically incorporates all other influences, including the influence of those "other parties" identified above. Figure 4 illustrates this point.

Thus, we believe that the allocation of natural gas savings, as established in partnership agreements with other parties before program launch, should apply to gross program savings. Specifically, it should

8 The *Filing Guidelines* actually state that an explanation is needed when the natural gas utilities' allocated share of natural gas savings in the partnership agreement is "more than 20% of" (i.e., more than one-fifth of) the "percentage of total dollars spent" (p. 22). We believe this is not consistent with the example provided in a footnote of the *Filing Guidelines*, but the interpretation in the text of the current document is consistent with that example.

come out of the assessed free-ridership portion of gross savings. If done after the application of free-ridership adjustments, then the utility is penalized twice for the effects of the same external influences.

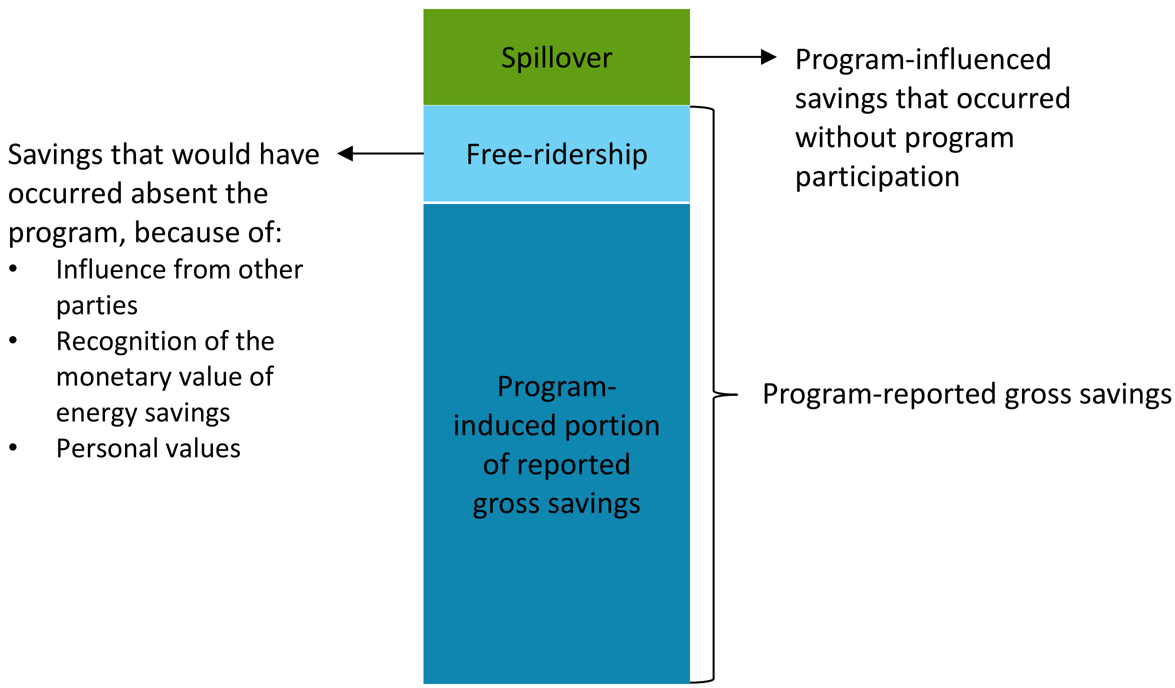


Figure 4: Components of Program Gross and Net Savings, Including Attribution to Other Parties

6. Conclusions and Recommendations

Regulators are rightly concerned about ensuring that energy efficiency programs not receive credit for energy savings that they did not cause. If there are no checks in place to ensure accurate assessment of attribution of savings, then program designers and implementers may not get the feedback needed to adjust and fine-tune programs to deliver the most savings possible for the dollars spent.

Yet the preceding sections of this report identify many potential problems with the way that net savings assessment has been conducted. Particularly problematic are self-report methods, which are very common for their low cost and ease of administration. Respondent self-selection bias as well as several very well-documented psychological propensities can easily result in over-estimation of free-ridership. A lack of statistical precision can produce estimates that may change notably from year to year. Spillover often is not included in NTG ratios and, when it is, it may very likely be under-estimated. Moreover, while the inclusion of spillover generally would increase the accuracy of a NTG estimate, it *decreases* the *precision* of NTG estimates because the separate estimates of free-ridership and spillover each contribute to the variance of the combined estimate. On top of all of the above – or perhaps, largely as a result of it – applying variable and unpredictable NTG adjustments retroactively can lead to conflict and litigation from dissatisfied shareholders.

Even apart from the above considerations, some evaluators (e.g., Violette et al. 2015) have concluded that the value of annual NTG research may not justify the cost. This conclusion applies even to self-report methods, which are probably the least expensive primary research methods to implement.

What, then, is the alternative to conducting program-year-specific primary NTG research? Based on our foregoing review and analysis, we offer the following recommendations to OEB and the natural gas utilities:

- **Develop a negotiated (also called "deemed" and "stipulated") NTG value.** This value should be based on a range of inputs, including a review of researched NTG values from similar programs in comparable jurisdictions that account for free-ridership and spillover, at a minimum, but also market effects if possible. Assessment of applicable NTG values from multiple studies should not treat all inputs equally but should follow a meta-analytic approach, which includes reviewing the study quality, assessing study heterogeneity, and developing a pooled estimate of variability based on the variabilities reported in the studies. The pooled estimate is a better representation of what the true estimate is in the population and it can provide insight into variability around NTG that are important to consider when determining what the value should be. Part of reviewing study quality should include assessing efforts taken to reduce the self-report biases identified in section 3. Other inputs to the negotiated NTG value should include structured expert judgment and any available market data or macroeconomic analyses. In developing the negotiated value, it may be valuable to employ a "value of information" approach, such as described by Violette et al. (2015).
- **Allocate any savings to parties other than the program only from the free-ridership portion of gross savings.** By definition, free-ridership represents the program-claimed savings that would have occurred without program assistance, which must include savings attributable to other parties. Allocating savings net of free-ridership to other parties doubly penalizes the program.

As noted in the body of this report, establishing a negotiated NTG value does not preclude doing NTG research, as such research may be valuable for program planning and implementation as well as to inform periodic adjustments to the negotiated NTG value. We recommend that OEB and the natural gas utilities observe the following when NTG research is conducted:

- **Always include spillover and, if feasible, market effects assessments.** As documented in the body of this report, failure to account for these factors will underestimate NTG.
- **If using self-report, employ methods to reduce the bias toward high free-ridership.** Energy Trust of Oregon, with input from Research Into Action, Inc., developed an approach to free-ridership assessment that attempts to control for the high-free-ridership bias of other self-report methods in addition to reducing customer fatigue (see Bliss, McClaren, Folks, and Kociolek, 2015; Roy and Bliss 2012). This alternative approach balances the counterfactual assessment with a component that assesses the influence of the various program interventions, which typically produces a lower free-ridership estimate than the counterfactual (PWP and Evergreen Economics 2017).
- **Assess free-ridership as close as possible to project implementation.** The longer the time that has elapsed between the implementation of the project and the assessment of the decision-making that went into the project, the less salient the external influences (including the program influence) will be to the program participant and the more likely that participant will be affected by the biases toward free-ridership responses.
- **Use multiple methods and triangulate the NTG estimate.** The use of multiple methods, such as surveys of contractors as well as program participants, is now generally regarded as best practice among energy efficiency experts (Kushler et al. 2014; PWP and Evergreen Economics 2017).

Following the above recommendations may allow the natural gas utilities to continue offering large C&I customers in Ontario opportunities to generate high energy savings through custom programs that may not otherwise be achievable.

Works Cited

Baumgartner, R. 2013. "Survey Design and Implementation Cross-Cutting Protocols for Estimating Gross Savings." In *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*, edited by T. Jayaweera, H. Haeri, pp. 12-1–12-25. Golden, CO: National Renewable Energy Laboratory (Subcontract Report NREL/SR-7A30-53827). Accessed at: www.nrel.gov/docs/fy13osti/53827.pdf.

Beggan, J. K. 1992. "On the Social Nature of Nonsocial Perception: The Mere Ownership Effect." *Journal of Personality and Social Psychology*, 62, 229-237.

Bliss, R., N. Sage, and D. Diebel. 2017. "Not all Spillover is the Same – So Don't Treat it That Way!" In *Proceedings of the International Energy Program Evaluation Conference*. Madison, WI: IEPEC.

Bliss, R., M. McClaren, J. Folks, and E. Kociolek. 2015. "From Real-time to Over-time: Developing a Four-year Perspective on an Energy Efficiency Portfolio." In *Proceedings of the International Energy Program Evaluation Conference.* Madison, WI: IEPEC.

Cadmus (The Cadmus Group, Inc.). 2017. *Assessment of Energy and Capacity Savings Potential in Iowa.* Prepared for The Iowa Utility Association. February 28, 2012. Accessed August 2, 2017 at: http://www.iowautility.org/wp-content/uploads/2016/05/Final-IUA-Joint-Assessment-Report-Vol.-1.pdf.

Dickerson, C. A., R. Thibodeau, E. Aronson, and D. Miller. 1992. "Using Cognitive Dissonance to Encourage Water Conservation." *Journal of Applied Social Psychology*, 22(11), 841-854.

Festinger, L. 1957. *A Theory of Cognitive Dissonance*. Evanston. IL: Row, Peterson.

Frisch, D. and J. Baron. 1988. "Ambiguity and Rationality." *Journal of Behavioral Decision Making*. 1 (3): 149–157. doi:10.1002/bdm.3960010303.

GDS Associates, Inc., Nexant, Inc., and Mondre Energy, Inc. 2014. *Act 129 Statewide Evaluator Final Annual Report, Phase I: June 1, 2009 – May 31, 2013*. Prepared for Pennsylvania Public Utility Commission. http://www.puc.pa.gov/pcdocs/1274547.pdf.

GDS Associates, Inc., Nexant, Inc., Research Into Action, Inc., and Apex Analytics, LLC. 2015. *Act 129 Statewide Evaluator Annual Report, Program Year 5: June 1, 2013 – May 31, 2014*. Prepared for Pennsylvania Public Utility Commission. http://www.puc.pa.gov/Electric/pdf/Act129/SWE_PY5-Final_Annual_Report.pdf.

GDS Associates, Inc., Nexant, Inc., Research Into Action, Inc., and Apex Analytics, LLC. 2016. *Act 129 Statewide Evaluator Annual Report, Program Year 6: June 1, 2014 – May 31, 2015*. Prepared for Pennsylvania Public Utility Commission. http://www.puc.pa.gov/Electric/pdf/Act129/SWE_PY6-Final_Annual_Report.pdf.

GDS Associates, Inc., Research Into Action, Inc., and Apex Analytics LLC. 2017. *Act 129 Statewide Evaluator Phase II Final Report*. Prepared for Pennsylvania Public Utility Commission. http://www.puc.pa.gov/Electric/pdf/Act129/Act129-SWE_PhaseII_FinalAR.pdf.

Haeri, H. and M.S. Khawaja. 2012. "The Trouble with Freeriders." *Public Utilities Fortnightly*. March 2012. Accessed August 2017. https://www.fortnightly.com/fortnightly/2012/03/trouble-reeriders.

Hall, N., L. Megdal, P. Jacobs, S. Waterbury, R. Wright, P. Chernick, K. Keating, S. Barata, E. Vine, S. Nadel, M. Kushler, R. Prahl, and J. Reed. 2004. *The California Evaluation Framework*. Retrieved August 8, 2017 from www.calmac.org.

Keating, K. M. 2009. "Freeridership Borscht: Don't Salt the Soup." In *Proceedings of the International Energy Program Evaluation Conference*. Madison, WI: IEPEC.

Khawaja, M.S., H. Haeri, and B. Hedman. 2014. "Finding Common Ground on Energy Efficiency." *Public Utilities Fortnightly*. December 2014.

Kunda, Z. 1987. "Motivated Inference: Self-serving Generation and Evaluation of Causal Theories." *Journal of Personality and Social Psychology*, 53(4), 636.

Kushler, M., S. Nowak, and P. Witte. 2012. "A National Survey of State Policies and Practices for the Evaluation of Ratepayer-Funded Energy Efficiency Programs." Prepared for ACEEE (American Council for an Energy-Efficient Economy). http://ilsagfiles.org/SAG_files/Subcommittees/IPA-TRC_Subcommittee/6-16-2015_Meeting/NEBs_Sources/ACEEE_2012%20report.pdf

Kushler, M., S. Nowak, and P. Witte. 2014. "Examining the Net Savings Issue: A National Survey of State Policies and Practices in the Evaluation of Ratepayer-Funded Energy Efficiency Programs." Prepared for ACEEE (American Council for an Energy-Efficient Economy). http://aceee.org/sites/default/files/publications/researchreports/u1401.pdf

Leidos (Leidos Engineering, LLC). 2014. Annual Report, Program Year 2014. Prepared for the Hawaii Public Utilities Commission by Hawaii Energy. December 11, 2015. https://hawaiienergy.com/images/resources/ProgramYear2014_AnnualReport.pdf

Messenger, M., R. Bharvirkar, B. Golemboski, C.A. Goldman, and S.R. Schiller. 2010. "Review of Evaluation, Measurement, and Verification Approaches Used to Estimate the Loan Impacts and Effectiveness of Energy Efficiency Programs." Lawrence Berkeley National Laboratory: Environmental Energy Technologies Division. LBNL-3277E. https://emp.lbl.gov/sites/all/files/lbnl-3277e.pdf

Miller, D.T. and M. Ross. 1975. "Self-serving Biases in the Attribution of Causality: Fact or Fiction?" *Psychological Bulletin*, 82, 213 – 225.

Ontario Energy Board 2015a. *Transition of Technical Evaluation Committee Activities to the OEB / EB-2015-0245.* Memorandum to Enbridge Distribution Inc., Union Gas Limited, Technical Evaluation Committee (TEC), and Evaluation Advisory Committee. March 4, 2016. https://www.oeb.ca/oeb/_Documents/EB-2014-0134/letter_TEC_EAC_Transition_20160304.pdf.

Ontario Energy Board 2015b. *Ontario Natural Gas Technical Evaluation Committee 2015 3rd Quarter Report.* March 10, 2016. https://www.oeb.ca/documents/TEC/Quarterly%20TEC%20Reports/Q3%202015%20TEC%20Report_March%20TEC.pdf.

Ontario Energy Board. 2014a. *Report of the Board: Demand Side Management Framework for Natural Gas Distributors (2015-2020) / EB-2014-0134*. December 22, 2014. https://www.oeb.ca/sites/default/files/uploads/Report_Demand_Side_Management_Framework_20141222.pdf.

Ontario Energy Board 2014b. *Filing Guidelines to the Demand Side Management Framework for Natural Gas Distributors (2015-2020) / EB-2014-0134*. December 22, 2014. https://www.oeb.ca/oeb/_Documents/EB-2014-0134/Filing_Guidelines_to_the_DSM_Framework_20141222.pdf.

Peters, J.S. and M. McRae. 2008. "Freeridership Measurement Is Out of Sync with Program Logic or, We've Got the Structure Built, but What's Its Foundation?" *In Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings.* Monterey, CA: American Council for an Energy-Efficient Economy. http://aceee.org/files/proceedings/2008/data/papers/5_491.pdf

Peters, J., R. Bliss, P. Degens, and S. Castor. 2010. "Faster Feedback: Real-Time Estimation of Free Riders." Poster presented at the Energy Efficiency Global Forum, Washington, DC, May 10-12.

PWP, Inc. and Evergreen Economics. 2017. *Current Methods in Free Ridership and Spillover Policy and Estimation*. Prepared for Energy Trust of Oregon. https://www.energytrust.org/wp-content/uploads/2017/07/FR_Spillover_170206.pdf.

Regional Technical Forum (RTF). 2010. *Methods for Evaluating Behavior-Based Energy Conservation Programs in the Pacific Northwest*. Retrieved August 3, 2017 from https://nwcouncil.app.box.com/s/qgolil20tzoqadphg2612jyxat8fp4jn.

Ridge, R., P. Willems, J. Fagan, and K. Randazzo. 2009. "The Origins of the Misunderstood and Occasionally Maligned Self-Report Approach to Estimating Net-to-Gross Ratio." In *Proceedings of the International Energy Program Evaluation Conference*. Madison, WI: IEPEC.

Roy, L., M. Johnson, P. Burns, J. Smith, W. Hodgson, C. Seruto, N. Benton, G. Sidorov, S. Hees, R. Bliss, and P. Schwarz. 2016. *Impact Evaluation of Washington Natural Gas 2014-2015 Energy Efficiency Programs*. Prepared for Avista Utilities by Nexant, Inc. with Research Into Action, Inc. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiBp9jf9OTVAhXJ5yYKHX-kD04QFggmMAA&url=https%3A%2F%2Fwww.utc.wa.gov%2F_layouts%2F15%2FCasesPublicWebsite%2FGetDocument.ashx%3FdocID%3D32%26year%3D2013%26docketNumber%3D132046&usg=AFQjCNF5GydBM7ln1-7sTchqghRnRH3v6w

Roy, L. and R. Bliss. 2012. "Findings, Challenges and Lessons Learned from the Evaluation of ARRA-Funded Programs for the Colorado Governors Energy Office." Paper presented at the AESP Annual Conference, February 6-9.

Saxonis, W. 2007. "Free Ridership and Spillover: A Regulatory Dilemma." In *Proceedings of the International Energy Program Evaluation Conference*. Madison, WI: IEPEC.

SBW (SBW Consulting, Inc.), Research Into Action, Inc., New Horizon Technologies, Inc., and Ridge & Associates. 2013. *Impact and Process Evaluation of NorthWestern Energy 2007-2011 Demand Side Management Programs.* Prepared for NorthWestern Energy. http://puc.sd.gov/commission/dockets/gaselectric/2012/ge12-001/vol1.pdf

Schwarz, N. 2007. "Retrospective and Concurrent Self-reports: The Rationale for Real-time Data Capture." In A. A. Stone, S. S. Shiffman, A. Atienza, & L. Nebeling (Eds.), *The Science of Real-time Data Capture: Self-reports in Health Research* (pp. 11-26). New York: Oxford University Press.

SEEAction (State & Local Energy Efficiency Action Network). 2012. *Energy Efficiency Program Impact Evaluation Guide*. Prepared by Steven R. Schiller, Schiller Consulting, Inc., https://www4.eere.energy.gov/seeaction/system/files/documents/emv_ee_program_impact_guide_0.pdf.

Stone, J., A. W. Wiegand, J. Cooper, and E. Aronson. 1997. "When Exemplification Fails: Hypocrisy and the Motive for Self-integrity." *Journal of Personality and Social Psychology*, 72(1), 54.

Skumatz, L. and E. Vine. 2010. "A National Review of Best Practices and Issues in Attribution and Net-to-Gross: Results of the SERA/CIEE White Paper." Prepared for ACEEE (American Council for an Energy-Efficient Economy). http://aceee.org/files/proceedings/2010/data/papers/2078.pdf.

Synapse Energy Economics, Inc. 2015. *Ontario Gas Demand Side Management 2016-2020 Plan Review*. Prepared for Ontario Energy Board Staff. http://www.synapse-energy.com/sites/default/files/Ontario_Gas_DSM_Plan_Review_15-065.pdf.

TecMarket (The TecMarket Works Team). 2006. *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. Prepared for the California Public Utilities Commission. http://www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=5212

Tetra Tech, Inc. 2011. *Cross-Cutting C&I Free-Ridership and Spillover Methodology Study Final Report*. Prepared for Massachusetts Program Administrators. http://ma-eeac.org/wordpress/wp-content/uploads/Massachusetts-PAs-Cross-Cutting-CI-Free-ridership-and-Spillover-Methodology-Study.pdf.

Vine, E., N. Hall, K. Keating, M. Kushler, and R. Prahl. 2010. "Emerging Issues in the Evaluation of Energy-Efficiency Programs: The US Experience." Published online November 26, 2010 at Springerlink.com.

Violette, D., K. Seiden, J. Erickson, M. Podolefsky, J. McMillan, S. Robinson, S. Dimetrosky, and N. Lieb. 2015. *Final Report: Iowa Energy-Efficiency Net-to-Gross Report.* Prepared for Iowa Utility Association and the Oversight Committee. September 2, 2015. http://www.iowautility.org/wp-content/uploads/2016/04/Appendix-B-Iowa-Net-to-Gross-Final-Report-11-25-2015.pdf.

Violette, D. and L. Agapay-Read. 2016. "Programme Attribution Analysis: Use of Self-Reports and Triangulation." Presented at the *International Energy Policy & Programme Evaluation Conference*, Amsterdam, June 7-9. http://www.ieppec.org/wp-content/uploads/2016/05/Paper-Violette-1.pdf.

Violette, D. and P. Rathbun. 2014. "Estimating Net Savings: Common Practices." In *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*, edited by T. Jayaweera, H. Haeri, pp. 23-1–23-74. Golden, CO: National Renewable Energy Laboratory (Subcontract Report NREL/SR-7A30-53827). Accessed at: www.nrel.gov/docs/fy13osti/53827.pdf.

Northeast Energy Efficiency Partnerships (NEEP) 2016. “Gross and Net Savings Principles and Guidance.” Prepared for Northeast Energy Efficiency Partnerships by Navigant Consulting and Tetra Tech, April 2016. http://www.neep.org/gross-and-net-savings-principles-and-guidance.

Chapter 21: Estimating Net Savings – Common Practices

The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures

Created as part of subcontract with period of performance September 2011 – September 2016

This version supersedes the version originally published in September 2014. The content in this version has been updated.

Daniel M. Violette, Ph.D.
Navigant
Boulder, Colorado

Pamela Rathbun
Tetra Tech
Madison, Wisconsin

NREL Technical Monitor: Charles Kurnik

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Subcontract Report
NREL/SR-7A40-68578
October 2017

Contract No. DE-AC36-08GO28308

NATIONAL RENEWABLE ENERGY LABORATORY

Chapter 21: Estimating Net Savings – Common Practices

The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures

Created as part of subcontract with period of performance September 2011 – September 2016

This version supersedes the version originally published in September 2014. The content in this version has been updated.

Daniel M. Violette, Ph.D.
Navigant
Boulder, Colorado

Pamela Rathbun
Tetra Tech
Madison, Wisconsin

NREL Technical Monitor: Charles Kurnik

Prepared under Subcontract No. LGJ-1-11965-01

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Subcontract Report
NREL/SR-7A40-68578
October 2017

Contract No. DE-AC36-08GO28308

This publication was reproduced from the best available copy submitted by the subcontractor.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at SciTech Connect http:/www.osti.gov/scitech

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
OSTI http://www.osti.gov
Phone: 865.576.8401
Fax: 865.576.5728
Email: reports@osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
NTIS http://www.ntis.gov
Phone: 800.553.6847 or 703.605.6000
Fax: 703.605.6900
Email: orders@ntis.gov

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.

Disclaimer

These methods, processes, or best practices ("Practices") are provided by the National Renewable Energy Laboratory ("NREL"), which is operated by the Alliance for Sustainable Energy LLC ("Alliance") for the U.S. Department of Energy (the "DOE").

It is recognized that disclosure of these Practices is provided under the following conditions and warnings: (1) these Practices have been prepared for reference purposes only; (2) these Practices consist of or are based on estimates or assumptions made on a best-efforts basis, based upon present expectations; and (3) these Practices were prepared with existing information and are subject to change without notice.

The user understands that DOE/NREL/ALLIANCE are not obligated to provide the user with any support, consulting, training or assistance of any kind with regard to the use of the Practices or to provide the user with any updates, revisions or new versions thereof. DOE, NREL, and ALLIANCE do not guarantee or endorse any results generated by use of the Practices, and user is entirely responsible for the results and any reliance on the results or the Practices in general.

USER AGREES TO INDEMNIFY DOE/NREL/ALLIANCE AND ITS SUBSIDIARIES, AFFILIATES, OFFICERS, AGENTS, AND EMPLOYEES AGAINST ANY CLAIM OR DEMAND, INCLUDING REASONABLE ATTORNEYS' FEES, RELATED TO USER'S USE OF THE PRACTICES. THE PRACTICES ARE PROVIDED BY DOE/NREL/ALLIANCE "AS IS," AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING BUT NOT LIMITED TO THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL DOE/NREL/ALLIANCE BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER, INCLUDING BUT NOT LIMITED TO CLAIMS ASSOCIATED WITH THE LOSS OF PROFITS, THAT MAY RESULT FROM AN ACTION IN CONTRACT, NEGLIGENCE OR OTHER TORTIOUS CLAIM THAT ARISES OUT OF OR IN CONNECTION WITH THE ACCESS, USE OR PERFORMANCE OF THE PRACTICES.

Preface

This document was developed for the U.S. Department of Energy Uniform Methods Project (UMP). The UMP provides model protocols for determining energy and demand savings that result from specific energy-efficiency measures implemented through state and utility programs. In most cases, the measure protocols are based on a particular option identified by the International Performance Verification and Measurement Protocol; however, this work provides a more detailed approach to implementing that option. Each chapter is written by technical experts in collaboration with their peers, reviewed by industry experts, and subject to public review and comment. The protocols are updated on an as-needed basis.

The UMP protocols can be used by utilities, program administrators, public utility commissions, evaluators, and other stakeholders for both program planning and evaluation.

To learn more about the UMP, visit the website, https://energy.gov/eere/about-us/ump-home, or download the UMP introduction document at http://www.nrel.gov/docs/fy17osti/68557.pdf.

Acknowledgments

The chapter authors wish to thank and acknowledge the Uniform Methods Project Steering Committee and Net-to-Gross Technical Advisory Group members for their contributions to the original Net Savings chapter and/or this updated version of the chapter. The following people offered valuable input to the development of this chapter by providing subject-related materials, in-depth discussion, and careful review of draft versions:

- Michael Li, U.S. Department of Energy
- Chuck Kurnik, National Renewable Energy Laboratory
- Michael Rufo, Itron, Inc.
- Hossein Haeri, M. Sami Khawaja, and Arlis Reynolds, The Cadmus Group, Inc.
- Tom Eckman, Northwest Power Planning Council
- Elizabeth Titus, Northeast Energy Efficiency Partnerships
- Steve Schiller, Schiller Consulting
- Rick Ridge, Ridge & Associates
- Ralph Prahl, Ralph Prahl & Associates
- Jane Peters and Marjorie McRae, Research Into Action, Inc.
- Frank Stern, Ken Seiden, and Jeff Erickson, Navigant
- Lynn Hoefgen, NMR Group, Inc.
- Nick Hall, TecMarket Works
- Miriam Goldberg, DNV GL
- Peter Miller, Natural Resources Defense Council
- Dick Spellman, GDS Associates, Inc.
- Tim Xu, Pacific Gas and Electric Company
- Rafael Friedmann, Pacific Gas and Electric Company
- Ryan Firestone, Consultant to the Regional Technical Forum (RTF), Northwest Power and Conservation Council (NWCouncil)
- Christian Douglass, Consultant to the Regional Technical Forum (RTF), Northwest Power and Conservation Council (NWCouncil)
- Bill Provencher, University of Wisconsin (University of Wisconsin, Agricultural and Applied Economics)
- Teri Lutz, formerly of Tetra Tech, made substantive contributions across the entire chapter.

Suggested Citation

Violette, Daniel M.; Rathbun, Pamela. (2017). *Chapter 21: Estimating Net Savings – Common Practices: Methods for Determining Energy-Efficiency Savings for Specific Measures*. Golden, CO; National Renewable Energy Laboratory. NREL/SR-7A40-68578. http://www.nrel.gov/docs/fy17osti/68578.pdf

Acronyms

C&I	commercial and industrial
CFL	compact fluorescent lamp
DiD	difference-in-differences
EE	energy efficiency
EM&V	evaluation, measurement, and verification
FR	free-ridership
FOE	Focus on Energy
HER	Home Energy Report
IOU	investor-owned utility
ISP	Industry Standard Practice
kWh	kilowatt-hours
LFER	linear fixed-effects regression
MCM	macroconsumption metric
ME	market effects
NPSO	nonparticipant spillover
NTG	net-to-gross
NW Council	Northwest Power and Conservation Council
PSO	participant spillover
RCT	randomized control trial
RDD	regression discontinuity design
RED	random encouragement design
RTF	Regional Technical Forum
SRA	self-report approach
TRM	technical reference manual
UMP	Uniform Methods Project

Protocol Updates

The original version of this protocol was published in September 2014.

This chapter has been updated to incorporate the following revisions:

- Modified the definitions of net and gross savings
- Reorganized the chapter slightly by:
    - Dividing the section on experimental design into two separate sections—one focusing on approaches that use random assignment (e.g., randomized control trials) and a second addressing quasi-experimental design approaches
    - Adjusting the order in which methods are presented to improve the logical flow of the chapter
- Expanded the discussion of survey methods based on recent developments in the literature
- Updated the Common Practice Baseline section with examples of how they have been set.

Table of Contents

List of Tables

1 Estimating Net Energy Savings

This chapter focuses on the methods used to estimate net energy savings in evaluation, measurement, and verification (EM&V) studies for energy efficiency (EE) programs. The chapter provides a definition of net savings, which remains an unsettled topic both within the EE evaluation community and across the broader public policy evaluation community, particularly in the context of attribution of savings to a program. The chapter differs from the measure-specific Uniform Methods Project (UMP) chapters in both its approach and work product. Unlike other UMP resources that provide recommended protocols for determining gross energy savings, this chapter describes and compares the current industry practices for determining net energy savings but does not prescribe methods.

Readers should treat this chapter as a resource document that provides state-of-the-art information about common practices for determining net energy savings. The selection and description of methods are based on the results of research by EM&V experts. The chapter describes the common methods and the approaches that are receiving attention in the evaluation community and discusses how net savings values are used for reporting and for energy-system planning.

The determination of net savings is an issue in EE programs that are funded publicly or through utility-customer resources. For these programs, the most direct contribution of net savings evaluation studies is to provide decision-makers the information they need to make good EE investments. Program goals, scale, funding sources, and the specific audience for the evaluation effort can influence the methods used, the aspects of the evaluation that are emphasized, the depth of analysis, and the way the results are presented.

Estimating net savings is central to many EE evaluation efforts and is broad in scope. It requires the determination of baselines (i.e., the counterfactual) and savings levels across many types of programs. The intent of this document is to present information on the tradeoffs in the various methods for calculating net savings that will help policy-makers, regulators, and program administrators decide which are best to apply.

The references section at the end of this chapter includes cited articles that address the presented methods in greater depth than the scope of this chapter allows.

2 Universality of the Net Impacts Challenge

Investment decisions result in allocating resources to achieve objectives. Regardless of the type of investment, once made, it is difficult to assess what would have happened absent that decision. This is the essence of evaluation: "What are the impacts of that investment decision?" These are termed *net impacts*, or *attributable impacts*. To address net impacts, a baseline is needed that represents what would have happened in the absence of the investment. This baseline is also called the *counterfactual scenario*.[1]

The broader literature on evaluation reveals a parallel between issues arising from estimating the net impacts of EE investments and estimating the effects of other types of investments made in either the private or the public sector. Examples include:

- Healthcare: What would the health effects have been without an investment in water fluoridation?
- Tax subsidies for economic development: Would the project—or a variant of the project—have proceeded without a subsidy?
- Education subsidies: What would happen if school lunch programs were not subsidized or if low-interest loans for higher education were not offered?
- Military expenditures: What would have happened without an investment in a specific military program or technology?

Across industries and applications, program evaluators grapple with how to appropriately approximate the counterfactual scenario and determine impacts that are attributable to the investment being analyzed (Cook et al. 2010). [2]

[1] As discussed in the "Considering Resource Constraints" section of the UMP *Chapter 1: Introduction*, small utilities (as defined under U.S. Small Business Administration regulations) may face additional constraints in undertaking this protocol. Therefore, alternative methodologies should be considered for such utilities.
[2] Some evaluators also view net savings estimation as an assessment of causality. This chapter uses the term *attribution* rather than *causality*, as it is more descriptive of the problem discussed, whereas causality has a wider range of interpretations that extends to metaphysics.

3 Defining Gross and Net Savings for Practical Evaluation

This section defines key terms related to estimating net savings and summarizes various uses of their measurement in the industry. It also describes many issues evaluators face when estimating net savings in the context of developing an appropriate baseline against which program accomplishments are compared to estimate net impacts.

3.1 Definition of Gross and Net Savings

The following definitions of gross and net savings are used in this chapter:

- **Net savings:** The difference in energy consumption *with the program in place* versus what consumption would have been *without the program in place*.
- **Gross savings:** The difference in energy consumption *with the energy-efficiency measures promoted by the program in place* versus what consumption would have been *without those measures in place*.

3.2 Definitions of Factors Used in Net Savings Calculations

The factors most often considered in net savings calculations are free-ridership, spillover (both participant and nonparticipant), and market effects. The definitions of these factors shown in Section 3.2.1 and Section 3.2.2 are consistent with those contained in the Energy Efficiency Program Impact Evaluation Guide (SEE Action 2012b). Not all net estimation methods require the explicit estimation of these factors, but they are useful considerations when assessing how completely different estimation methods address net savings in the context of attribution.

3.2.1 Free-Ridership

Free-ridership is the program savings attributable to free-riders (program participants who would have implemented a program measure or practice in the absence of the program). There are three types of free-riders:

- **Total free-riders:** Participants who would have completely replicated the program measure(s) or practice(s) on their own and at the same time in the absence of the program.
- **Partial free-riders:** Participants who would have partially replicated the program measure(s) or practice(s) by implementing a lesser quantity or lower efficiency level.
- **Deferred free-riders:** Participants who would have completely or partially replicated the program measure(s) or practice(s) at a time after the program timeframe.

3.2.2 Spillover

Spillover refers to additional reductions in energy consumption or demand that are due to program influences beyond those directly associated with program participation. As a result, these savings may not be recorded in the program tracking system and credited to the program. There are generally two types of spillover:

- **Participant spillover:** This represents the additional energy savings that are achieved when a program participant—as a result of the program's influence—installs EE measures or practices *outside* the efficiency program after having participated.

  Evaluators have further defined the broad category of participant spillover into the following subcategories:

  - *Inside spillover:* Occurs when participants take additional program-induced actions at the project site
  - *Outside spillover:* Occurs when program participants initiate actions that reduce energy use at sites that are not participating in the program
  - *Like spillover:* Refers to program-induced actions participants make outside the program that are of the same type as those made through the program (at the project site or other sites)
  - *Unlike spillover:* Refers to EE actions participants make outside the program that are unlike program actions (at the project site or other sites) but that are influenced in some way by the program

- **Nonparticipant spillover:** This represents the additional energy savings that are achieved when a nonparticipant implements EE measures or practices as a result of the program's influence (for example, through exposure to the program) but is not accounted for in program savings.

3.2.3 Market Effects

Market effects refer to "a change in the structure of a market or the behavior of participants in a market that is reflective of an increase in the adoption of energy efficiency products, services, or practices and is causally related to market intervention(s)" (Eto et al. 1996). For example, programs can influence design professionals, vendors, and the market (through product availability, practices, and prices), as well as influence product or practice acceptance and customer expectations. All these influences may induce consumers to adopt EE measures or actions (Sebold et al. 2001).[3]

Some experts suggest that market effects can be "best viewed as spillover savings that reflect significant program-induced changes in the structure or functioning of energy efficiency markets." Prahl et al. (2013) also suggest that market transformation is a subset of market effects

[3] When assessing EE policies in a broad context, it should be acknowledged that some participants identified as free-riders in a current program might not have had the opportunity to adopt the EE measure or service were it not for the effects on the market from previous EE program efforts. These efforts may have contributed to that measure or service being available to customers in the current year. The importance of this issue to evaluation depends on the parameters of the evaluation. Most evaluations focus on set time periods spanning 1–3 years. Factors that are included are based on the incremental actions taken as a result of the EE program year being evaluated and the current state of the EE market. Actions taken that resulted from EE efforts in preceding years represent sunk costs and are not incremental to the current program being evaluated. However, this may be an important consideration in a broader policy assessment examining the overall trend in the adoption of EE measures and services across a longer time period. Market effects of previous years' programs may not have been fully accounted for, and this can be a consideration in the broader policy context. However, for assessing the impacts of a given EE program for a given year, these effects from past programs are not generally considered. This is discussed in more detail in Section 4.3.

(as the substantive and long-lasting effects). This view implies that market effects are a subset of spillover. Although spillover and market effects are related, the methods used to quantify these two factors generally differ. Therefore, this chapter addresses them separately.

3.2.4 Net Savings Equations

Evaluators use different factors to estimate net savings for various programs and jurisdictions depending on how a jurisdiction views equity and responsibility (NMR Group, Inc. and Research Into Action 2010). For example, some jurisdictions include only free-ridership in the calculation of net savings; others include both free-ridership and spillover. Some jurisdictions estimate net savings without measuring free-ridership or spillover (market-level estimates of net savings). Messenger et al. (2010) also discuss differences across jurisdictions in the reporting of gross and net savings.

A practitioner who is trying to develop methods to estimate values for these factors will find the definitions provided in this section useful. However, the evaluator must work with the information available, which starts with the tracking system.[4] Evaluators typically view the data in the tracking system as the initial estimate of gross savings. Because free-ridership, spillover, and market effects are untracked values, evaluators should estimate or account for them outside of the program tracking system.[5] A practical way to understand these values is to consider spillover and market effects as savings that are attributable to the program, but that are not included in the program tracking system. Free-ridership represents savings included in the program tracking system that are not attributable to the program.

To estimate net savings, the evaluator first estimates free-ridership, spillover, and market effects, then makes appropriate adjustments to the values in the tracking database (or validated tracking database) as illustrated in Equation 1.[6]

---

[4] The definitions for *free-ridership*, *spillover*, and *market effects* should be integrated with (1) how the utility tracks actual program participation data; and (2) how the utility records information about expected program impacts in the program tracking system. In general, the initial gross savings estimate (in terms of expected energy savings by participant or measure) comes from the tracking system. These data may include "deemed values" negotiated by the stakeholders. These deemed values may include factors that lower the savings of a measure, based on assessments of current practice, codes and standards, and other factors that may directly or indirectly influence how the estimated gross savings are adjusted to estimate net savings. It is important to understand how the gross savings are estimated by project and by participant. In fact, the first recommendation of NMR Group Inc. and Research Into Action (2010) is that the Northeast Region needs a process leading to the development of a consistent definition of *adjusted gross savings*.

[5] Direct estimation methods are available to address free-ridership, spillover, and market effects without estimating each separately. This chapter addresses randomized control trials, quasi-experimental designs, and common practice baselines, each of which essentially is used to adjust the savings estimates in the program tracking system.

[6] A *validated tracking database* is simply a reviewed program tracking database. A review of the tracking database can determine obvious errors, whether adjustments can make the claimed (*ex ante*) savings entries more accurate, and whether any deemed savings values include adjustments that account for net savings factors (for example, an adjusted baseline that captures market trends). The validated tracking system then contains the most accurate information on claimed savings for each participating site or project. The benefits of improved information in the tracking system are discussed by Violette et al. (1993).

Equation 1. Net Savings Including Free-Ridership, Spillover, and Market Effects

Net Savings = Gross Savings – FR + SO + ME not already captured by SO

Where:

FR = free-ridership savings

SO = spillover savings

ME = market effects savings not already captured by SO

In much of the literature, the program evaluation approach involves a net-to-gross (NTG) ratio for which free-ridership, spillover, and market effects are expressed as a ratio to gross savings (Equation 2). These widely used ratios work well for some types of evaluation efforts (for example, survey-based estimations). The term is almost synonymous with estimating net savings and is commonly defined as the ratio of NTG savings for the sample. The population gross savings is then multiplied by the NTG ratio to estimate population net savings.

Equation 2. Net-to-Gross Ratio

NTG Ratio = 1 – FR ratio + SO ratio + ME ratio (where the denominator in each ratio is the gross savings)

When using the NTG ratio defined by specific free-ridership, spillover, and market-effect factors (or ratios), evaluators use Equation 3 to calculate net savings:

Equation 3. Net Savings Calculation Using the Net-to-Gross Ratio

Net Savings = NTG Ratio * Gross Savings

These definitions are essentially standard in the evaluation literature;[7] however, a given jurisdiction may decide not to include free-ridership, spillover, or market effects to estimate net savings. For example, evaluators almost always include free-ridership, but, because of policy choices made in a jurisdiction, most do not always fully consider spillover and market effects (see NMR Group, Inc. and Research Into Action 2010; NEEP 2012). Most evaluators agree that spillover and market effects exist and have positive values, but determining the magnitudes of these factors can be difficult. Increasingly, the trend is to include estimates of spillover in net savings evaluations. The inclusion of market effects is also increasing, but to a lesser degree than spillover. Methods are available to address spillover and market effects and, because there is no debate about whether they exist, these factors should be addressed or considered when estimating net savings. For market effects, evaluators must first conduct a basic assessment as to whether market effects are likely, as well as consider the cost and value of a study. It is important to

[7] Other factors (sometimes called *net impact factors*) are generally considered as adjustments to gross impact estimates. These include rebound, snapback, and persistence of savings. Violette (2013) addresses these factors. As with other NTG factors, evaluations do not treat net impact factors consistently in gross impact calculations, and do not consistently adjust program gross impacts to calculate to a final net impacts number.

know the potential sizes of spillover and market effects for a given program or portfolio so appropriate policy decisions can be made about EE investments.

3.3 Uses of Net Savings Estimates in the Energy Efficiency Industry

Many regulatory jurisdictions discuss the appropriate use of net savings estimates. This is due in part to: (1) the cost of the studies to produce these estimates and (2) a perceived lack of confidence in the resulting estimates.[8] However, evaluators and regulators recognize the advantages of consistently measuring net savings over time as a key metric for program performance (Fagan et al. 2009).

Evaluators generally agree that net savings research can be useful for (SEE Action 2012a, 2012b):[9]

- Gaining a better understanding of how the market responds to the program and using that information to modify the program design (including eligibility and target marketing and incentive levels).
- Gleaning insight into market transformation over time by tracking net savings across program years and determining the extent to which free-ridership and spillover rates have changed over time. This insight might be used to define and implement a program exit strategy.
- Informing resource supply and procurement plans, which requires an understanding of the relationship between efficiency levels embedded in base-case load forecasts and the additional net reductions from programs.
- Assessing the degree to which programs effect a reduction in energy use and demand (net savings is one program success measure that should be assessed).

With respect to the last bullet, Schiller (SEE Action 2012b, pp. 2-5) also discusses the importance of consistently measuring savings across evaluation efforts and having consistent evaluation objectives. For example, evaluators in different jurisdictions assess the achievement of goals and targets as measures of overall EE program performance using different measures of savings: gross savings, net savings, or a combination of the two. There are also differences across jurisdictions where the measure of EE program success is used for calculating financial incentives. There are arguments for basing financial incentives on net savings, as well as arguments for basing incentives on gross savings or a combination of the two.[10]

---

[8] Several experienced evaluators indicated in comments on earlier drafts of this chapter that in their experience, the required level of confidence and precision for estimates of net impacts within the EE field is generally greater than that used in other fields faced with similar types of questions and tradeoffs. The authors generally agree with this observation, but no meta-study comparing target levels of confidence and precision for EE program evaluation with similar evaluations in other fields has been conducted.

[9] Other methods that can and should be used to inform program design and understand market response include process evaluations and market assessments.

[10] As more jurisdictions begin to consider the delivery of EE programs as a business process that requires an investment of resources, they are considering the return on investment (more commonly termed *incentives*), which is typically coupled with performance targets. Jurisdictions can base targets on reaching a certain level of gross savings

A recent NEEP document (NEEP 2016) provides six core principles to inform states' decision regarding their applications of gross and net savings based on policy goals. These six principles are: (1) establish a common understanding, (2) align methods and use with policies, (3) address the value of information from evaluation, (4) apply the concept of symmetry,[11] (5) ensure transparency, and (6) acknowledge multiple views across stakeholders.

3.4 The Net Savings Estimation Challenge—Establishing the Baseline

This chapter discusses estimation methods that rely on the development of a baseline (the assumed counterfactual scenario). This baseline is used to measure the net impacts of a program. If evaluators could identify a "perfect baseline" (i.e., a counterfactual scenario that exactly represents what would have happened if the EE program had not been offered), most of the issues associated with estimating net impacts would not arise.

The evaluator is faced with the challenge of identifying a method that produces a baseline that best represents the counterfactual scenario—in other words, what the participant group (and the market) would have done in the absence of the program.[12] To understand and defend the selection of a method for estimating net savings, the evaluator should consider the implicit and explicit assumptions used for the baseline. For example, when considering the use of nonparticipants as a candidate baseline, the evaluator needs to account for issues that pertain to the similarity, or matching, of the program participants with customers that may comprise the nonparticipant comparison group. The evaluator should also account for any effects the program might have had on the comparison group that may influence the program net savings.

Self-selection can be viewed as a baseline issue that arises when a program is voluntary and participants select themselves into the program, suggesting the potential for systematic differences between program participants and nonparticipants. This issue is not unique to EE evaluations and arises in any policy or program assessment involving self-selection. Specifically, the assumption in this case is that the self-selected participants are those who would have taken more conservation actions than the general nonparticipant comparison group.[13]

---

or on achieving a certain level of net savings—each has pros and cons. A gross savings target may provide a clearer incentive structure for the program administrator, and there is generally less controversy over whether the target is achieved. The fact that incentives are usually based on a calculation of shared benefits, where the predominant share of benefits goes to ratepayers, creates an equitable incentive structure: the program administrator receives fewer benefits and even if attributed (net) savings are lower than expected, the ratepayers still receive most of the benefits. For example, under an 80%–20% split of the benefits (80% of benefits are realized by ratepayers and 20% by the administrator), having attributed savings reduced by 50% still implies that 70% of the benefits go to ratepayers. See Rufo (2009) for other views on aligning incentives with the outputs of program evaluation.

[11] Symmetry refers to recognizing all the components of net savings – both positive and negative influences, and recognizing the impact on the net savings estimate when not all components are included.

[12] Agnew and Goldberg (2017) provide a number of choices for selecting comparison groups for use in billing analyses. This chapter also discusses using regression analysis as a tool for making appropriate comparisons and arriving at alternative net savings values.

[13] In this case, the nonparticipant baseline does not fully correct for free-riders, resulting in estimated net savings that are biased upward. Other self-selection factors could cause the participant and nonparticipant groups to behave differently. For example, if participants need the financial assistance to make the investment and nonparticipants do not need the rebate to take EE actions, the baseline comparison group might take more EE actions than the participant group in the absence of the program. In this case, a nonparticipant baseline would produce estimated net

Free-ridership reduces net program savings in this example case, but other variants of self-selection might increase net savings when a participant group is compared to a nonparticipant baseline. For example, if the customers who self-select into the program need the financial incentives to justify the EE investment, an adjustment for self-selection might increase overall net savings.

Spillover can also be viewed as a baseline issue. For example, nonparticipant spillover can occur when the energy consumption of the comparison group of nonparticipants is not indicative of what the energy consumption for this group would have been in the absence of the program. In this case, the comparison group is *contaminated:* the program affected the behavior of those in the comparison group.

This section discussed issues related to establishing an appropriate baseline as an approximation of the counterfactual scenario. Understanding that free-ridership, spillover, and market effects can be viewed as baseline issues can help the evaluator focus on the factors that are most important to selecting an appropriate method.[14] In many applications, selecting the baseline is a core issue in choosing an appropriate estimation method. When presenting the net savings results of a program, the evaluator should include a description of the baseline and the assumptions implicit in the estimation method.

---

savings that are biased downward and appropriately correcting for this self-selection effect would increase the estimated net savings. The authors have observed that often there is an assumption that addressing self-selection will always lower estimated net savings by reducing bias caused by free-riders, but this is not always the case.

[14] Self-selection, free-ridership, and spillover issues are not unique to EE evaluation—they are common in other settings as well. Consider a business decision made to produce net benefits, such as downsizing. Might self-selection be important to address in assessing this business initiative? Employees who have the best experience and are the most confident in their ability to land new jobs might (if able) self-select into the downsizing option. Might there be some free-riders if the downsizing effort includes personnel who were planning to leave anyway? Also, there might be spillover impacts from the downsizing program where having workers leave reduces the productivity of employees who remain. Although self-selection, free-ridership, and spillover pose challenges for EE evaluation, these same issues often have to be addressed in evaluating investment decisions in other fields and contexts.

4 Methods for Net Savings Estimation

This section discusses methods for estimating net savings, as well as some of the advantages and challenges associated with each. Evaluators use a variety of methods, some of which address free-ridership and/or spillover (for example, self-report surveys); others focus on market effects (for example, structured judgment approaches or historical tracing). The methods addressed in this section are:

- Randomized control trials (RCTs) and options for randomized approaches
- Quasi-experimental designs including matching
- Survey-based approaches
- Market sales data analyses
- Structured expert judgment approaches
- Deemed or stipulated NTG ratios
- Historical tracing (or case study) method
- Common practice baseline approaches
- Top-down evaluations (or macroeconomic models).

Table 1 lists methods that are applicable for estimating free-ridership, spillover, and market effects. This table indicates the general applicability of the methods. The following sections review the specific applications, caveats, limitations, and other key information in greater detail to explain how to assess the methods for each net savings component.

Table 1. Applicability of Approaches for Estimating Net Savings Factors

Method	Free-Ridership	Spillover	Market Effects
Randomized control trials and options for randomized approaches	Controls for free-riders[a]	Controls for participant spillover[b]	Not generally used
Quasi-experimental designs including matching	Controls for free-riders[c]	Controls for participant spillover	Not generally used; however, if the design includes observations over multiple years, then some market effects can be captured
Survey-based approaches	Is applicable	Is applicable	In conjunction with structured expert judgment
Market sales data analysis	Is applicable	Is applicable	Is applicable
Structured expert judgment[d]	Is applicable	Is applicable	Is applicable
Deemed or stipulated NTG ratios	Is applicable	Is applicable	Not generally used
Historical tracing	Is applicable	Is applicable	Is applicable
Common practice baseline methods	Is applicable	Not applicable[e]	Not applicable
Top-down evaluations	Assess the overall attributable change in energy use, so no separate adjustment is needed for free-ridership, spillover, and market effects[f]		

[a] Does not provide a direct estimate of free-ridership, but rather controls for free-riders through experimental design.
[b] Does not estimate spillover, but rather controls for participant spillover through experimental design. A separate study of control group members is required to address nonparticipant spillover if it is expected to be significant and affect the net impacts.
[c] Like RCTs, these designs do not provide a direct estimate of free-ridership, but self-selection bias can still be present. Unlike RCTs, the choice aspect of opt-in EE programs may not be fully addressed, unless additional methods are applied.
[d] This approach is applicable only if the experts are knowledgeable about the specific market being studied.
[e] Spillover could arguably be addressed through surveys of participants and nonparticipants, but this is not generally viewed as being part of the common practice baseline method, and the use of surveys would make this more similar to survey-based estimation methods discussed in Section 4.3.
[f] However, depending on the details of the analysis, these elements may not be fully captured.

More discussion on applicable methods for different types of residential and commercial programs and the pros and cons of these different methods can be found in a 2014 supplemental guidance document prepared for NEEP (NEEP 2014).

4.1 Randomized Controlled Trials and Options for Related Randomized Approaches

This section discusses random controlled trials (RCTs) and options for related random assignment approaches. RCTs represent the ideal approach and produce net savings accounting for free-ridership, participant spillover and avoid the problem of self-selection by addressing the potential choice-based biases by random assignment. However, RCT approaches may not always be possible. When an RCT is not possible, the quasi-experimental designs in Section 4.2 can be used as alternatives. RCTs can be difficult to set up and more applications are seen with pilot programs. However, RCTs are increasingly being used to evaluate behavioral programs, information programs, and pricing programs designed to increase efficiency. Generally, most RCT applications have been in the residential sector where large numbers of customers (both participants and controls) are available to the researcher. Even if a pure RCT is not possible, other approaches can be used to take advantage of random assignment. These other approaches, including random encouragement designs (RED) and random recruit or deny or structural criteria to avoid opt-in biases such regression discontinuity designs (RDD), have seen their applications to EE evaluation increase.

4.1.1 Randomized Control Trials (RCT)

An RCT design is ideal for assessing the net impacts of a program—particularly the free-ridership and short-term spillover components. If the RCT is short term (that is, 1 year or less), it may not be able to address any longer-term spillover, and addressing spillover and market effects may require additional data collection efforts for each year of the study.

For the RCT, the study population is defined first, then consumers from the study population are randomly assigned to either a treatment group (participants in the EE program) or to a control group that does not receive the treatment (nonparticipants). Random assignment is a key feature of this method. By using random probability to assign consumers to either the treatment or the control group, the influence of observable differences between the two groups is eliminated (for example, location of home, age of home, and appliance stock). Unobservable differences are also eliminated (for example, attitudes toward energy use, expectations about future energy prices, and expertise of household members in areas that might induce participation) (NMR Group, Inc. and Research Into Action [2010]; SEE Action [2012a, 2012b]). This method, when implemented properly, can provide a near-perfect baseline that results in reliable net savings estimates addressing free-riders and self-selection.

The net savings calculations are relatively straightforward when an RCT is designed properly. The literature generally covers three methods for calculating net savings:

1. **Use a simple post-period comparison to determine the differences in energy use between the control and treatment groups after participation in the program.** For example, if participating households are using 15,000 kilowatt hours (kWh) on average and the control households are using 17,000 kWh, the net savings estimate is 2,000 kWh.
2. **Use a difference-in-differences (DiD) approach to compare the change in energy use for the two groups between the pre- and post-participation periods**. For example, assume participants used 17,500 kWh prior to program participation and 15,000 after participation, for a difference of 2,500 kWh between the pre- and post-periods. Assume also that the well-matched control group has similar pre-period energy use (approximately 17,500 kWh), but the group's post-period energy use is 17,000 kWh (that is, slightly lower, possibly because of weather), for a difference of 500 kWh. Applying the DiD method results in an estimated savings of 2,000 kWh (the 2,500 kWh change for participants minus the 500 kWh change for nonparticipants).
3. **Use a linear fixed-effects regression (LFER) approach**, where the regression model identifies the effects of the program by comparing pre- and post-program billing data for the treatment group to the billing data for the control group. A key feature of the LFER

approach is the addition of a customer-specific intercept term that captures customer-specific effects on electricity use that do not change over time, including those that are unobservable. Examples include the square footage of a residence, the number of occupants, and thermostat settings (see Provencher and Glinsmann [2013] for an example and additional discussion of the LFER method).[15]

Even if randomizing the treatment and control groups, an evaluator may use a method other than the simple post-period comparison to be as thorough as possible and use all the available data to develop the estimate. The DiD method tracks trends over time, and the fixed-effects component of the LFER adds an extra control for the differences between consumers that are constant during the period being examined. All three methods generate unbiased estimates, as randomization ensures no systematic differences between the treatment and control groups in the drivers of energy use, so the three methods would be expected to generate similar, but not necessarily identical, results.

The RCT approach is simple in concept, but may be more difficult to implement given available data, timing, program design, and program implementation issues. It is becoming standard practice for evaluators to use statistical methods to test whether the allocation of customers between the treatment group and the control group is consistent with what would be expected from a random assignment of consumers to the treatment and control groups.[16] For billing data, this type of analysis often involves comparing the means of the two groups with respect to demographic variables (if available) and monthly energy use in the pre-program year. For example, if the differences in means for the two groups fall outside a 90% confidence bound for more than 2 months of the pre-program year, there is cause for concern that assignment to the two groups is not random. (See an example of an application of this test for consistency with RCT expectations in Provencher and Glinsmann [2013] and other tests in Stuart [2010]). If this is the case, it is worth examining how the random assignment was conducted to ensure no inadvertent elements of the process are affecting assignment to the treatment and control groups. The goal of this testing is to determine if non-random factors are affecting the assignments, not to keep repeating the random selection process until the samples fit an ideal profile. If several characteristics are compared, it is not unusual to have some that are "significantly" different between the two groups. Regression analysis helps to mitigate these effects.

The RCT approach to estimating program impacts reflects the "intent to treat" effect. Generally, it is not appropriate to drop customers after the random assignment, though the consequences of doing so vary. For example, questions may arise about what to do with consumers who opt out.

[15] A number of the methods discussed in this chapter use regression approaches. Some are fairly simplistic; others are quite sophisticated, requiring expertise in econometrics. Each section provides citations to applied studies, many of which describe the econometric techniques employed. For example, Stuart (2010) lists econometric software and routines that can be useful in matching. Also, Agnew, and Goldberg (2017) discuss regression models in more detail but provide a limited set of literature references. SEE Action (2012a) recommends Greene (2011) as a useful reference on regression techniques. Wooldridge (2010) focuses on cross-section and panel data models that are often used in evaluation. Kennedy (2008) and Angrist and Pischke (2008) are useful supplements to any econometrics textbook.

[16] Even with random assignment, it is important to apply best practices in the design of analysis including stratification both to reduce standard errors (increase precision) and help ensure representativeness of the sample drawn.

Consider, for instance, a program involving Home Energy Reports (HERs), in which program administrators send energy use reports by mail. This program was designed to generate energy savings by providing residential consumers information about their energy use and energy conservation. Some percentage of consumers will opt out of the program. They should remain in the analysis because the similar set of control consumers who would have opted out of the program could not be identified if they were to receive the report. Also, on average, these consumers might have different energy use than the other control consumers, causing the reported impact to be biased if the treatment group is adjusted to remove the opt-out consumers. At the other extreme, HERs might not be deliverable because of observable address characteristics. If this same address characteristic can also be identified for control consumers, the estimate of program impacts after eliminating treatment and control consumers with this characteristic is, strictly speaking, an unbiased estimate of the effect of intent-to-treat *conditional* on the address characteristic. These examples are meant to show that careful analysis is needed in the application of all methods, including RCTs. In addition, Duflo et al. (2007) caution that excessive investigations of subgroups not specified *ex ante* constitute a form of data mining that should be avoided. The case discussed above where address characteristics are available for the treatment and control groups does not fall in this category, but this caution deserves emphasis.

To maintain an RCT over a period, evaluators must take care when working with the data across the treatment and control groups. For example, a behavioral program (such as HERs) may be rolled out to 20,000 high-use residential consumers in program year 1. In program year 2, an additional 20,000 consumers of all energy use classifications may enroll, and another 30,000 consumers may enroll in program year 3. Additionally, some consumers in program year 1 may have dropped out (requested to not receive the HERs).[17] Each of these sets of participating customers need to be appropriately considered in the RCT design and the appropriate assignment of customers to be used as controls.

Issues inevitably arise about the consumer energy use data. Researchers have used the following criteria, among others, as indicators of problems with consumer billing data:

- Having fewer than 11 or more than 13 bills during a program year
- Having fewer than 11 or more than 13 bills during the pre-program year
- Energy consumption outside a reasonable range (that is, an outlier observation with average daily consumption that is lower than the 1st percentile or higher than the 99th percentile)
- Observations with fewer than 20 or more than 40 days in the billing cycle.

Agnew and Goldberg (2017) also discuss issues with consumer energy use data and program data in residential settings. Even programs that have operated for several years are likely to have issues. Using the HERs example, this could include consumer records that are missing the date when the first report was sent or entries in consumer records that indicate issues with that observation.

[17] This is not an unusual problem in the utility industry. Utilities have for many years addressed similar issues in maintaining random customer samples for load research.

After addressing data issues, the evaluator probably still has a good RCT, unless many consumers are affected by these data issues or consumers are disproportionately affected across the participant and control groups. Mort (2017) presents additional criteria that can cause sites to be excluded and suggestions about what to do if the number of removed sites exceeds 5%.

The ability to disseminate information to large groups of consumers has led to an increase in RCTs in EE evaluation.[18] In general, these RCT-based evaluations have focused on residential behavior-based EE programs such as HERs programs. These programs lend themselves to random trials in that they: (1) provide information only; (2) can be implemented for relatively homogeneous consumers at the same time; and (3) allow for an RCT design. These characteristics, however, are not generally present for many large-scale EE programs that tend to account for many of the EE portfolio savings.

In summary, the RCT approach is the most accurate method for estimating net impacts. The RCT controls for free-riders and near-term participant spillover—two important factors. To the extent that the program affects the control group, nonparticipant spillover is not addressed. This effect is likely to be small over the short run in most behavioral programs. If nonparticipant spillover is large, net impacts will be underestimated because there are nonparticipants who were affected by the program, and the baseline will be inaccurate. To appropriately address this issue, the evaluator would need to conduct a separate study of control group members to address nonparticipant spillover. Because market effects are longer term spillover effects, they would likely not be included in any RCT net savings approach that spans just a few years.

Although the RCT method can produce an accurate baseline when constructed correctly, it is not always possible to apply an RCT to evaluations of EE programs for a variety of reasons. RCT generally requires planning in advance of program implementation. As pointed out in Chapter 8 (Agnew and Goldberg 2017) of these protocols, "…evaluation concerns have been less likely to drive program planning." Also, an RCT approach may involve denying or delaying participation for a subset of the eligible and willing population. In some cases, the random assignment may result in providing services to consumers who either do not want them or may not use them (see Table 2 for pros and cons of RCTs).

Other characteristics of programs that can make an RCT difficult to implement include:

- Programs that require significant investments, such as a commercial and industrial (C&I) major retrofit program in which the expenditures are in the tens of thousands of dollars. Typically, these programs are opt-in, and random assignments within an eligible study population might include consumers who either do not need the equipment or services or do not want to make that investment. Programs that involve relatively large investments

[18] Evaluations of HERs programs that used RCTs include Sacramento Municipal Utility District (2011), Puget Sound Energy (2012), AEP (2012), PG&E (2013), Commonwealth Edison (2012), and Pacific Gas & Electric (2013). Some ongoing evaluations use RCT methods for HERs programs, and will produce additional practical information on RCT applications. Another useful study, but one focused on evaluating pricing programs, which used an RCT design is the Sacramento Municipal Utility District (2013). This study assesses different pricing structures in the residential sector; however, the methods used are good examples of what can also be applied in EE evaluations in an RCT context.

in measures and services across the residential and C&I sectors may not amenable to an RCT design.

- C&I programs often have participants that are more heterogeneous than is the case for residential programs which would require large samples of both treatment and control groups than may be available. In some cases, a few very large customers can be relatively unique within a utility service area or region with few similar consumers who might be appropriate candidates for a control group.
- To achieve savings targets, programs may be rolled out over an entire year, with consumers opting in every month. As a result, consumers self-select into the participant group, which is unknown until after 1 year of the program implementation. Evaluators can more easily apply RCT to programs with a common start date for many participants (for example, HERs programs). There are ways to address this, but this adds somewhat to the complexity of the design. The random recruit and deny design discussed below can be used to addressing rolling program roll-outs.

4.1.2 Other Forms of Randomization and Approaches for Minimizing Opt-in Selection Bias

Two other approaches incorporate random assignment to help address the choice-based bias of opt-in programs—random encouragement design (RED) and random recruit deny/delay approaches. Another approach that can be used to minimize opt-in selection bias is the regression discontinuity method (RDD). RDD does not incorporate randomization. Instead, it looks for cutoffs or discontinuities in participation that can be used to construct two eligible participant groups.

4.1.2.1 Random Encouragement Design

Random encouragement design (RED) is also applicable to the types of data available for EE program evaluation. Rather than being randomly assigned to a treatment or control group, customers are randomly assigned to receive supplemental encouragement to participate in the program (e.g., a letter informing a random set of customers about a rebate), or not to be so encouraged. RED involves taking a randomly selected group of participants to receive extra encouragement, which typically takes the form of additional information or incentives. A successful encouragement design allows the effects of the intervention and encouragement to be estimated (Diamond and Haninmueller 2007; McKinzie 2009[19]). In this case, there may be an EE program for which all consumers can decide to opt in such as a residential audit program or a commercial audit or controls programs. A group of randomly selected consumers is then provided extra encouragement in terms of information and/or financial incentives. This randomization can ameliorate the effects of self-selection.[20]

[19] In a position statement closely related to what EE program evaluators face, McKenzie states that “Rigorous impact evaluations, which compare the outcomes of a program or policy against an explicit counterfactual of what would have happened without the program or policy, are one of the most important tools that can be used along with appropriate economic theory for understanding ‘what works.’ Despite this, until recently impact evaluations have been rare, especially outside the areas of health and education.”

[20] The underlying estimation concept in RED is explained by the U.S. Department of Energy (2010): “In RED, researchers indirectly manipulate program participation using an encouragement ‘instrument’ so as to generate the

The RED design provides the average net savings per participant for those who participate because of the encouragement but otherwise would not. This is not necessarily the same as the net savings for the original program without extra encouragement. In particular, we would expect free-ridership to be lower among those who need extra encouragement. Thus, the RED might be expected to overstate net savings for the original program if free-ridership is present but would still provide useful information.

Fowlie and Wolfram (2009) outline an application of RED to a residential weatherization program and address the design of the study and Fowlie, Greenstone and Wolfram (2015) apply this design to a low-income program.[21] They point out that:

> *REDs are particularly useful when:*
>
> - *Randomization of access or mandatory participation is not practical or desirable.*
> - *There is no need to ration available services (that is, demand does not exceed supply).*
> - *The effects of both participation and outreach are of interest to policy makers.*

Rather than randomize over the intervention, the encouragement to participate is randomly manipulated. This allows the effect of the encouragement to produce exogenous variation in program participation, which can help identify the effect of the program on participants (U.S. Department of Energy 2010).

Evaluators should take certain practical issues into account in any research design, and RED is no exception. The sample sizes needed for an RED study are typically larger than for a pure RCT, and groups receiving the encouragement need to show different participation rates.[22] Evaluators should consider this research design when estimating net savings, as it aligns well with many standard EE program implementation plans. The random variation is designed not by excluding participants but simply by providing enhanced information and/or incentives offered to the selected consumers. Ongoing research work using RED should provide useful information for practitioners and the EE evaluation community. RED is growing with most applications focused on residential programs.

4.1.2.2 Random Recruit Deny/Delay

Finally, another approach that can be used to construct randomized treatment and control groups is the random recruit and deny/delay design. In this case, the timing of the treatment is randomly assigned. Customers are recruited with the understanding that they will randomly be assigned to receive the program offering immediately or later. The control group is thus a randomly selected set of customers that have opted-in but receive the treatment later. This is an effective way of

---

exogenous variation in program participation that is so essential for causal inference. This exogenous variation can then be used to identify the effect of the program on those households whose participation was contingent upon the encouragement." Other useful references to RED are Bradlow (1998) and West (2008).

[21] Fowlie et al. (2015) find limited energy savings from the weatherization assistance program that was evaluated. This find was challenged by Hogan (2015).

[22] This can be one of the challenges in the design of an RED approach. The design of the encouragement given to a random sample of participants must be effective; that is, produce higher acceptance rates than for the balance of the participant group.

ensuring that the control and treatment groups are well matched, but it may not fully address other types of selection. If it is a pilot program that is being evaluated, the savings impacts may be accurate for the pilot participants, but it may be difficult to extrapolate these impacts to a broader set of customers. There are two other issues: (1) The fact that a customer did opt-in to a program but had their participation delayed may, in itself, change their behavior (e.g., they may not take actions they would have taken in the absence of the program as they are expecting to receive the benefits of participation in the near future); and, (2) Some customers may drop out of the research pool if they learn their participation will be delayed and that they will be part of a control group. An example of this research design is Xcel Energy (2016).

4.1.2.3 Regression Discontinuity Design

SEE Action evaluation guides (2012a, 2012b) discuss the regression discontinuity design (RDD). This method is becoming more widely used and is applicable to programs where a cutoff point or other discontinuity separates otherwise likely program participants into two groups. This approach examines the impacts of a program by using a cutoff value that puts consumers into or out of the program through a design that does not involve their selecting themselves into the program or choosing not to participate. As a result, this approach addresses the self-selection issue.[23] By comparing observations lying closely on either side of a cutoff or threshold (i.e., the eligible and in-eligible cut off), the average treatment effect in environments where randomization is not possible can be estimated.[24] The underlying assumption in RDD is that assignment to participant and nonparticipant groups based on the eligibility cutoff produces groups that are otherwise similar. If this holds, those who just met the threshold for participating are comparable to those who just missed the cutoff and did not participate in the program and the difference in energy use between the two groups can reasonably be assumed to be the effect of program participation.

The SEE Action reports indicate that RDD can be a good candidate for yielding unbiased estimates of energy savings. The example used by SEE Action is based on an eligibility requirement for households to participate in a program. This requirement might be that a consumer whose energy consumption exceeds 900 kWh/month would be eligible to participate in a behavior-based efficiency program, while consumers who use less than 900 kWh/month would be ineligible. Thus, the group of households immediately below the usage cutoff level might be used as the comparison group.

For participating and nonparticipating households near the cutoff point of 900 kWh in monthly consumption, RDD is likely to be a good design. In the larger context, this RDD assumes that the program impact is constant across all ranges of the eligibility requirement variable (that is, the impact is the same for households at all levels of energy use). Evaluators should consider this

[23] In the recent years, there has been a strong movement toward focusing on the "identification" issue in evaluation; that is, the issue that in the absence of an RCT you do not really know if the error term in a regression is correlated with the explanatory variable of interest, so your estimate of the coefficient on that explanatory variable should be assumed to be biased in the absence of "sound" corrective action. A regression discontinuity design addresses this issue.
[24] The RDD has a history in evaluation dating back to the 1960s. This approach has been used to assess a wide variety of attribution analyses in the fields of education, health, and policy. Recently, this approach has been used more often. For a review of RDD see Imbens and Lemieux (2010).

assumption carefully for participating households that might consume much more than 900 kWh/month (for example, 2,000 kWh or more for some participants). Households with greater consumption may have greater opportunities for energy use reductions (although the change might be constant as a percentage). In this example, potential concerns about the consistency of program impacts across different levels of household energy use suggests an assessment of the quality of the resulting participant and control groups matched samples. Stuart (2010) has general guidance for assessing the quality of these designs.

Another discontinuity example is a time-based cutoff point. Because utilities often have annual budgets for certain programs, it is not uncommon for a program to exhaust its budget before the year is finished, sometimes within 6 months. In this case, a date-based cutoff is useful. Consumers who apply for the program after the enrollment cutoff date imposed by budget restrictions may be similar to the program participants accepted into the program during the first 6 months of the year. Also, both groups of consumers may have a more similar distribution of energy use per month (the focus of an impact assessment). This time-based cut-off approach is similar to using future participants as comparison groups discussed in UMP *Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol* (Agnew and Goldberg 2017).

4.1.3 Summary – RCTs and Related Randomization Approaches

Several types of approaches employing randomization are being used in evaluation. These include RCTs, REDs, and Recruit Deny/Delay. The RDD approach is a bit different as it takes advantage of a structural discontinuity (participation cut-off or threshold) to a treatment and control groups that are not affected by choice (i.e., opt-in programs).

The RCT approach is the most accurate method for estimating net impacts. The RCT controls for free-riders and near-term participant spillover—two important factors. To the extent that the program affects the control group, nonparticipant spillover is not addressed. This effect is likely to be small over the short run in most behavioral programs. If nonparticipant spillover is large, net impacts will be underestimated because there are nonparticipants who were affected by the program, and the baseline will be inaccurate. To appropriately address this issue, the evaluator would need to conduct a separate study of control group members to address nonparticipant spillover. Because market effects are longer term spillover effects, they would likely not be included in any RCT net savings approach that spans just a few years. These same caveats also apply to RED, Recruit Deny/Delay, and RDD approaches.

It is not possible to definitively determine whether the RED or Recruit Deny/Delay designs discussed above provide an appropriate comparison group. Fowlie and Wolfram (2009) point out that there have been studies comparing these designs to the ideal RCT. The finding is that randomized designs (either RED or RDD) improve on simple comparison approaches. RDD depends on the program having a cutoff point for participation that allows for random selection. RED may be a good fit with many EE programs that have many participants, but appropriate design in the types of information and incentives is required. Both RDD and RED depend on the assumption that the net savings of the isolated participants—those just under the threshold for RDD, and those who participate with only incremental encouragement for RED—is the same as the net savings for all participants.

Importantly, these methods should be considered in advance of program implementation to allow for the appropriate data, or the design of the information or incentives that will be offered to potential participants, to effectively implement these evaluation methods. It has always been important to consider evaluation when designing or revising EE programs, but for random assignment methods the evaluation method must be built into the program delivery.

Some of the pros and cons associated with these methods are presented in Table 2.

Table 2. Approaches using Random Assignment (RCTs, RED, Recruit Delay/Deny, and RDD) — Summary View of Pros and Cons

Pros	Random assignment reduces and limits bias in estimates Increases reliability and validity RCTs control for free-riders and participant spillover RCTs widely accepted in natural and social sciences as the highest standard of research designs RED, RDD, and Recruit Delay/Deny approaches also control for free-riders and participant spillover, but with additional assumptions regarding the appropriateness of each design.
Cons	Bias can result if random assignment occurs among volunteers or if the program drop-out rate differs by key characteristics Does not address nonparticipant spillover Equity/ethical concerns about assigning some ratepayers to a control group and not allowing them to participate in the program for a period of time May not be applicable to programs that involve large investments in measures and services Some C&I programs can have participants that are unique due to their size or industry, and there may be few control group candidates Needs to be planned as part of program implementation to allow for appropriate random assignment for RCT, RED, and Recruit Deny/Delay.

**This summary of pros and cons is not meant to replace the more detailed discussion in the text for guidance in application.*

4.2 Quasi-Experimental Designs

For most EE programs, either practical concerns or design factors will limit the use of RCT and other random assignment methods. In these situations, quasi-experimental designs are often a good option. Quasi-experimental designs are not unique to EE evaluations and are often used in evaluations of private and public investments. Stuart (2010) reviews the evolving research on matching and propensity scoring methods in quasi-experimental designs and states that such methods "… are gaining popularity in fields such as economics, epidemiology, medicine, and political science."[25,26]

[25] Stuart (2010) also provides a guide to software for matching, because software limitations have made it difficult to implement many of the more advanced matching methods. However, recent advances have made these methods more accessible. This section lists some of the major matching procedures available. A continuously updated version is also available at www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html. Common statistical software packages such as STATA, SAS, and R address most of the current matching approaches.

Quasi-experimental designs have some similarities to RCTs in terms of constructing comparison and treatment groups, except that random assignment is not possible. In a quasi-experimental design, consumers typically select themselves into the participant group, and the evaluation researcher must then develop the comparison group. To avoid confusion, quasi-experimental designs use the term *comparison group*, and RCT designs use the term *control group*. [27]

This section discusses two types of approaches to developing a comparison group within a quasi-experimental design–1) matching methods, and 2) panel data approaches. Matching methods use a measure of distance between two observations (e.g., customers) and can include Exact Distance, Mahalanobis Distance, and Propensity Scoring. Panel data approaches include structural regression modeling with a specific set of independent variables designed to address differences between the treatment and comparison groups (see the discussion of pooled regression in Agnew & Goldberg, 2017).

4.2.1 Matching Methods

Matching is broadly defined in the literature to be any method that aims to equate (or balance) the distribution of covariates in the treatment group and the comparison group.

The evaluator's goal is to select a comparison group that matches the participant group in terms of the actions that influence energy use. If done well, the only significant difference between the two groups will be participation in the program. Still, how well the comparison group actually matches the participant group will always be subject to some uncertainty, as there may be *unobservable* variables that affect energy use, the attribute of interest. Stuart (2010) defines the problem this way:

> *One of the key benefits of randomized experiments for estimating causal effects is that the treated and control groups are guaranteed to be only randomly different from one another on all background covariates, both observed and unobserved. Work on matching methods has examined how to replicate this as much as possible for observed covariates with observational (nonrandomized) data… While extensive time and effort [are] put into the careful design of randomized experiments, relatively little effort is put into the corresponding "design" of nonexperimental [quasi-experimental] studies. In fact, precisely because nonexperimental studies do not have the benefit of randomization, they require even more careful design.*

---

[26] Most attribution analyses assessing business decisions and public or private investments use quasi-experimental designs, as many practical factors result in the use of this method. As an extreme example, consider a study that is designed to assess the health effects of smoking. Would it be appropriate to select a study population of 9,000 18-year-olds and assign one third to a group that does not smoke, one third to a group that smokes a pack of cigarettes a day, and one third to a group that smokes a pack a day, but with some mitigating medications? Clearly, this type of RCT would pose ethical issues. As a result, natural quasi-experiments are used where smokers are matched with a comparison group of nonsmokers that is as representative as possible. The methods of matching on observable characteristics have become quite advanced in the past decade.

[27] Technically, quasi-experimental designs do not always include a nonparticipant comparison group. For example, the interrupted time-series design (Shadish et al. 2002) relies only on aggregate participant data over time and shows this method can help control for threats to internal validity; i.e., that the results of the study are appropriately estimated for the participating customers. External validity involves generalizing; i.e., the ability of the study results to be extrapolated to other groups of customers.

Stuart (2010) presents a good overview of the literature on matching and advantages of matching compared to regression models based on a set of explanatory variables. The recent evaluation literature, particularly for residential sector programs, shows the increasing use of matching. Recent approaches have focused on matching by energy use and energy use distributions across months and seasons. These matching methods can be simple or sophisticated, even when matching is confined to available energy use data (that is, no additional surveys of nonparticipants are conducted). Matching on energy use can be as simple as stratifying participants and nonparticipants by their energy consumption (season, year, or month) and then drawing nonparticipants to match the participants' distribution of energy use.

As discussed by Stuart (2010), the literature on matching based on energy use is expanding. Provencher and Glinsmann (2013) focus on a comparison of the distribution of energy across months and seasons. The analysis follows the approach advocated by Ho et al. (2007) and Stuart (2010). The procedure used by Provencher and Glinsmann (2013) first matches each participant household to a comparison household based on a minimum distance criterion—in this case, the minimum sum of squared deviations in monthly energy consumption for the 3 months of the specified season in the pre-program year.[28] In the second step, a regression model of the energy use of treatment customers and their matched controls, with covariates that include the matching variables, is used to identify the average treatment effect.

Matching methods tend to follow the literature reviewed by Stuart (2010). Stuart indicates that matching methods have four key steps, with the first three representing the "design" and the fourth the "analysis." These steps are:

1. Define closeness: the distance measure used to determine whether an individual is a good match for another.
2. Implement a matching method appropriate to the measure of closeness.
3. Assess the quality of the resulting matched samples (and perhaps iterate Step 1 and Step 2 until well-matched samples result).
4. Analyze the outcome and estimate the treatment effect, given the matching done in Step 3.

In Step 1, closeness is often defined as a minimum distance value as used in Provencher and Glinsmann. Another approach for identifying nonparticipants is "propensity scoring." The most common method used in propensity score estimation involves the estimation of a logistic regression. This model uses information about participants and nonparticipants to estimate a dependent variable assigned the value of 1 if that consumer is a participant or 0 if the consumer is a nonparticipant. This process allows for identification of nonparticipants who have similar

[28] In the program evaluation literature, matching often involves matching on variables with different metrics; for example, energy use and square footage of the household. These variables are normalized in the application of the distance criterion, usually using the full covariance matrix for the variables (the Mahalanobis metric). The original reference is Mahalanobis (1936) and the use of the metric is covered by Stuart (2010). One application, among many examples, is Feng (2006), which also includes the SAS® code for this method.

propensity scores to participants (that is, similar attributes between participants and nonparticipants). This approach has a long history in in the EE evaluation literature.[29,30]

The EE evaluation literature using matching methods (i.e., approaches that use a definition of closeness) has been expanding. Different types of applications that develop matching subject to constraints (e.g., a geographic constraint such as falling within a defined set of zip codes – See Navigant, 2016), and matching on hourly consumption rather than monthly data (PowerStream, 2016) are becoming more common. An application to a commercial sector pricing/thermostat program is found in Nexant (2017).

4.2.2 Panel-Data Models

Stuart (2010) states that alternatives to matching methods include adjusting for relevant covariates in a structured regression model. However, Stuart (2010) also points out that "matching methods should not be seen in conflict with regression adjustment and in fact the two methods are complementary and best used in combination."

One of the motivations for matching is to mitigate against model specification bias in the traditional structured regression panel-data model. Chapter 8 of the UMP (Agnew and Goldberg 2017) discusses consumption data analyses, including alternatives for constructing comparison groups. Also, the two SEE Action guides (2012a and 2012b) address matching. Matching methods include:

- **Participants as the comparison group:** SEE Action (2012b) states that among quasi-experimental approaches, "perhaps the most common [is] the 'pre-post' approach. With this approach, sites in the treatment group after they were enrolled in the program are compared with the same sites' historical energy use prior to program enrollment. In effect, this means that each site in the treatment group is its own nonrandom control group."

  By using the participant group as its own comparison group, the energy use of the participants during a period before they participated in the program is used as the

[29] The use of discrete choice methods to address self-selection in evaluations of EE programs has been presented in early evaluation handbooks. See Violette et al. (1991) and Oak Ridge National Laboratory (1991). More recently, Bodmann (2013) used a discrete choice model to develop an instrumental variable to address omitted variable bias. However, most of these applications occurred in the 1990s, probably because the development of a discrete choice model that has adequate predictive power requires large sample sizes, which make the surveys expensive to conduct. The discrete choice model needs to be able to predict customers who choose to participate and customers who choose not to participate with appropriate reliability. This approach thus requires both participant and nonparticipant surveys. This more advanced econometric topic is not dealt with in detail in this chapter; however, several reviewers believed it was important to provide references to these methods. Heckman (1979) originally developed the two-stage model for treating self-selection. These techniques are addressed both under instrumental variables and self-selection by Kennedy (2008), who states: "Selection is not well understood by practitioners. It rests fundamentally on the role of an unmeasured variable and so is similar to bias created by the omission of a relevant explanatory variable." (p. 286). An updated discussion of the Heckman models for self-selection, along with appropriate caveats, can be found in Guo and Fraser (2010). Note: a link to this chapter is provided in the References section. Guo and Fraser also show how the Heckman models relate to propensity scoring. Applications in the EE arena include Dubin and McFadden (1984), Goldberg and Kademan (1995), and Bodmann (2013), who used a discrete choice model to develop an instrumental variable to address omitted variable bias.

[30] Southern California Edison (2014) provides a recent behavioral impact application using propensity scoring.

comparison or baseline. A statistical consumption analysis is used that also includes factors that are expected to influence energy use and may vary across the pre-post time periods. Weather is the most obvious additional variable that should be controlled, but there may be other variables as well, such as economic factors if the periods cover a two-year period or longer. Agnew and Goldberg (2017) provide a useful set of algorithms for making weather adjustments.[31]

- **Nonparticipants as the comparison group:** The trend in the literature is to move away from the simple approach of using participants as their own comparison group in a time-series analysis and instead to develop cross-sectional time-series data that include data on participants and nonparticipants.

4.2.3 Summary of Quasi-Experimental Designs—Matching and Panel Data Regression Models

Randomized approaches may not always be possible to use. Quasi-experimental designs try to replicate designs that employ randomization using observational (nonrandomized) data. Matching as an evaluation method is rapidly expanding, particularly for residential programs. Panel data regression models can be used in conjunction with matching, or they can be used as stand-alone methods when data are available on relevant covariates and there is confidence in the appropriate structure for the models. Table 3lists some pros and cons with these approaches.

Table 3. Quasi-Experimental Designs—Summary View of Pros and Cons

Pros	Limits bias if a matched comparison group can be identified regarding the actions that influence energy use Unlike RCT, can be applied after program implementation Increases reliability and validity Partially controls for free-riders and participant spillover Widely accepted in natural and social sciences when random assignment cannot be used Matching may reduce concerns over model specification bias.
Cons	May be difficult to identify a matched comparison group if there are unobservable variables that affect energy use Does not address nonparticipant spillover Some C&I programs may have unique participants and few control group candidates Does not address self-selection bias without additional modeling, i.e., the estimation of a companion discrete choice participation model to address bias from choice-based participation in programs.

4.3 Survey-Based Approaches

This section describes the survey-based approach to collect NTG-related data and the analytic use of the data obtained. This approach can be a cost-effective, transparent, and flexible method for estimating NTG, and it has become one of the most often-used methods in EE net savings

[31]Other approaches can be used for weather normalization, particularly if the evaluator is interested in changes in monthly peak demand in addition to average monthly energy use. Additional weather normalization approaches are discussed by Eto (1988) and McMenamin (2008).

estimation. Consequently, it is important to understand good sample and survey design, and the strengths and weakness of these methods.

Surveys may target up to three types of respondents: (1) program participants, (2) program nonparticipants, and (3) market actors.[32] This section individually describes surveys with these three types of respondents; best practices recommend triangulating and using multiple survey approaches (for example, enhanced self-report) or multiple net savings estimation approaches.

The methods discussed in the preceding section provide estimates of net savings directly. That is, those approaches compare a participant group to either a random control group (as part of an RCT) or to a comparison group from a well-designed, quasi-experimental application, and these approaches do not require a separate effort to estimate free-ridership, spillover, or market effects.[33]

Survey-based approaches are used in evaluations that start with gross estimates, and then adjust for NTG factors. Surveys can be a cost-efficient means to estimate NTG factors, but they are not without issues, as discussed in the following subsections. Baumgartner (2013) also discusses many of the issues involved in using surveys to estimate NTG.

4.3.1 Program Participant Surveys

Survey-based methods for estimating net savings from program participants who are aware of the program incentives/services use questions about the program's influence on the participants' actions and decision-making. Participants answer a series of closed-ended and open-ended questions on these topics:

- Why they installed the program-eligible equipment.
- What they would have done in the absence of the program incentive and services.
- What further actions they took on their own because of their experiences with the program.

As noted by Baumgartner (2013), best practice survey design for attitudes and behavior measurement use multiple-item scales to better represent the construct. Because participant decision-making is complex, the survey should ask a carefully designed series of questions rather than a single question, as that could result in misleading findings. Refer to SEE Action (2012b), Megdal et al. (2009), Haeri and Khawaja (2012), and New York Department of Public Service (2013b) for discussions about the sequencing of a series of questions.

The primary benefits of a survey-based approach are:

- A survey approach can be less expensive than other approaches, particularly if the effort is combined with data collection activities that are already planned for process and impact evaluations.

[32] Note that a Delphi panel, which also uses surveys of a panel of experts, is discussed in Section 4.5 of this chapter.
[33] Market effects can be viewed as longer-term spillover effects; therefore, it is unlikely that any market effects are included in an RCT net savings approach spanning just a few years.

- The evaluator has the flexibility to tailor questions based on variations in program design or implementation methods.
- It can yield estimates of free-ridership and spillover without the need for a nonparticipant control group (NMR Group, Inc. and Research Into Action 2010). However, participant surveys capture only a subset of market effects,[34] a key piece of NTG.

Despite these benefits and the wide use of a survey-based self-report approach, significant concerns have been raised (Ridge et al. 2009; Peters and McRae 2008). The main concerns are:

- A potential bias related to respondents giving socially desirable answers.[35]
- The inability of consumers to know what they would have done in a hypothetical alternative situation, especially in current program designs that use multiple methods to influence behavior.
- The tendency of respondents to rationalize past decisions.
- A potential for arbitrariness in the scoring methods that translate responses into free-rider estimates.
- Consumers may fail to recognize the influence of the program on other parties who influenced their decisions. For example, a program having market effects may have influenced contractor practices, which in turn may have indirectly impacted the participants' (and nonparticipants') decisions.

Ridge et al. (2009) point out that, although these concerns are valid, they are widely acknowledged by social scientists who have worked on a variety of methods over the years to address them. It is also important to recognize that all methods have potential biases.[36] For example, market sales analysis,[37] which is based on objective sales data, can be biased if the market actors who provide data for the analysis operate differently from those not participating in the study or if the comparison area is systematically non-comparable.

In addition, Ridge et al. (2009) point out that it does not make sense to compare all self-report approaches equally, as some conform to best practice and others do not. Keating (2009) adds that many of the criticisms of the self-report approach can be alleviated through careful research design, sampling, survey timing, and wording of questions.

Baumgartner (2013) presents guidelines for selecting appropriate survey designs and recommends procedures for administering best practice surveys. The literature also contains

---

[34] Participant surveys can, in theory, capture end user market effects; for example, changes in end user awareness, knowledge, and efficiency-related procurement practices.

[35] Participants may also have a bias toward overstating program impacts because they want to retain incentives, although this has not been widely documented.

[36] This is, of course, the primary motivation for triangulation.

[37] Market sales analysis captures the total net effect of a program. Ideally, this method involves obtaining comprehensive pre- and post-market sales data in both the area of interest and in an appropriate comparison area and examining the change in the program area compared with the change in the non-program area (Tetra Tech et al. 2011).

several best practice elements for survey design, data collection, and analytic methods specific to estimating net savings (New York State Department of Public Service 2013; Tetra Tech et al. 2011). This literature notes the importance of making the entire process transparent so stakeholders can understand how each question and its response impacts the final estimate. Thus, the report should contain details of critical elements such as the question sequence, scoring algorithms, and the handling of inconsistent and/or missing data.

4.3.1.1 Survey Design Elements

Several design elements need to be considered when developing surveys. Best practices for choosing design elements include:

- Identify the key decision-maker(s) for the specific EE project. For downstream programs, a key decision-maker in the household or business is likely to be responsible for making the final decision, although they may assert that their vendor was the most influential in their decision. Although consumers ultimately decide what they will purchase, they may not be aware of the influence of the interventions for upstream programs where trade ally decisions are driving change (for example, original equipment manufacturers determine equipment EE levels and retailers determine what equipment to stock and market, or advertise as a result of upstream program incentives).
- Use setup or warmup questions to help the decision-maker(s) recall the sequence of past events and how these events affected their decision to adopt the measure.
- Use multiple questions to limit the potential for misunderstanding or the influence of individual anomalous responses.
- Use questions that rule out rival hypotheses for installing the efficient equipment.
- Test the questions for validity and reliability.
- Use consistency checks when conducting the survey to immediately clarify inconsistent responses.
- Use measure-specific questions to improve the respondent’s ability to provide concrete answers, and recognize that respondents may have different motivations for installing different measures.
- Use questions that capture partial efficiency improvements (accounting for savings above baseline but less than program eligible), quantity purchased, and timing of the purchase (where applicable for a measure) to estimate partial free-ridership.
- Use neutral language that does not lead the respondent to an expected answer.
- Use combinations of open- and close-ended questions to balance hearing from the end users in their own words and create an efficient, structured, and internally consistent dataset.

4.3.1.2 Data Collection Elements

Even when the survey design is effective, data collection should also follow best practices for collecting reliable information and calculating valid estimates. These practices include:

- Pretest the survey instrument to ensure that questions are understandable, skip patterns are correct, and the interview flows smoothly. The pretesting should use, when possible, cognitive interviewing techniques (Miller 2011).[38]
- Use techniques to minimize nonresponse bias, such as advance letters on utility or program administrator letterhead (the organization for which the participant will most likely associate the program) and multiple follow-ups over a number of weeks.
- Follow professional standards for conducting surveys, which include training and monitoring interviewers.[39]
- Determine the necessary expertise of the interviewer based on the complexity and value of the interview (for example, it is better for trained evaluation professionals rather than general telephone surveyors to address the largest, most complex projects in custom programs).
- Time the data collection so it occurs as soon as possible after a measure is installed, as this minimizes recall bias and provides timely feedback on program design. Recognize, however, that timely data collection for estimating free-ridership will underestimate participant spillover, as little time may have passed since program participation. Conducting a separate spillover survey later with these same participants can alleviate this. Having a separate survey will increase data collection costs, but may be warranted if spillover effects are likely to have occurred.
- Sample (or oversample) a census of the largest savers and, depending on program participation, sample end uses with few installations to ensure the measures are sufficiently represented in the survey sample.

4.3.1.3 Analytic Elements

In addition to discussing survey design and data collection elements, much of the literature discusses best practices for analysis such as:

- Treat acceleration of the installation of the EE measures appropriately to produce lifetime net savings rather than first-year net savings (this requires understanding the program's influence on the timing of the project).[40]
- Incorporate the influence of previous participation in the program.

[38] In cognitive interviews, respondents are asked to describe how and why they answered the question as they did. Miller (2011) notes that "through the interviewing process, various types of question response problems that would not normally be identified in a traditional survey interview, such as interpretive errors and recall accuracy, are uncovered." (p. 54).

[39] Data collections surveys can be conducted via telephone, the Web (including smartphones), postal mail, and in person. For large complex C&I projects, an energy engineer who is knowledgeable about the type of project and technology should conduct the interviews.

[40] Michael Rufo, Itron, notes that "A focus on program induced early replacement versus the effect on efficiency level is gaining attention in the evaluation field. In cases where there is early replacement, two net savings components may be needed to appropriately characterize overall net savings: (1) the early replacement period that uses an in-situ baseline; and, (2) the efficiency increment above minimum or standard practice at the end of the early adoption period (that is, one for the RUL (remaining useful life) period and one for the remainder of the EUL [effective useful life]."

- Establish *a priori* rules for treatment of missing/don't knows in the scoring algorithm.
- Weight the estimates by annual savings to account for the size of the savings impacts for each consumer.
- Sample, calculate, and report the precision[41] of the estimate for the design element of interest (measure, project type, or end use).
- Conduct sensitivity testing of the scoring algorithm.
- Define what the spillover measurement is and is not attempting to estimate and justify the use of an approach.
- Employ, where feasible, a preponderance of evidence (or triangulation of results) approach that uses data from multiple sources (see Itron, Inc. 2010), especially for large savers and complex decision-making cases. Potential data sources could include project file reviews, program staff and account manager interviews, vendor interviews, and observations from site visits.

The New York Department of Public Service (2012) developed additional guidelines specific to the estimation of spillover savings to address recurring methodological limitations that the New York Department of Public Service staff and its contractor team observed in the estimation of spillover in New York and the industry as a whole. Prahl et al. (2013) summarize this work and the critical decisions that evaluators must make before deciding whether and how to estimate spillover. That paper also discusses how the estimation of per-unit gross savings, estimation of program influence, and documentation of causal mechanisms varies for different levels of rigor.

4.3.2 Surveys of Program Nonparticipants

Self-report surveys with nonparticipants are commonly used to triangulate participant self-report responses and collect data for calculating nonparticipant spillover or market effects. These surveys help evaluators understand what EE actions nonparticipants have taken and whether they took those actions because of program influences (nonparticipant spillover). Conducting surveys with nonparticipants poses its own unique challenges:

- There is no record of the equipment purchase, and identifying a group of nonparticipants who have installed energy-efficient equipment on their own can be time consuming and costly.[42]
- Establishing causality entails estimating gross unit savings (often with limited evidence other than the consumer self-report) and establishing how the program may have influenced the consumer's decision. The consumer may not have been aware, for example, of the influence the program had on the equipment's availability or the market actor's stocking practices.

[41] The New York Department of Public Service (2013a) presents guidelines for calculating the relative precision of program net savings estimates for different types of estimates, including the NTG ratio based on the self-report method and for spillover savings. Additional discussion of sampling for evaluation can be found in Khawaja et al. (2013).
[42] One approach to mitigating the efficiency and cost of this is to use one nonparticipant survey that asks about a variety of program eligible measures and use the results across multiple programs.

4.3.3 Market Actor Surveys

When estimating net savings, it is important to consider all the points of program influence. In addition to targeting consumers, upstream and midstream programs often target program services and/or funding to market actors (such as contractors, auditors, and design specialists) with the goal of influencing their design, specification, recommendation, and installation practices. In upstream and midstream programs, consumers may not be aware of program influences on sales, stocking practices, or prices (discussed in the Appendix).[43] Thus, using only participant self-reports when estimating net savings is inappropriate. In these cases, evaluators use market actor self-report surveys to examine the effects of these upstream influences.

These market actor self-report surveys can be designed as qualitative in-depth interviews or as structured surveys with a statistically designed sample of contractors. The use and application of the data determine the format. For example, evaluators may use:

- Qualitative, open-ended data based on a small sample of market actors to contextualize market actors' practices (best used for triangulation purposes).
- Quantitative market actor data to calculate free-ridership and spillover rates specifically related to the practices of those market actors. The calculated rates can then be directly integrated with participant self-report results, triangulated with participant self-report results, and/or used as the sole source for free-ridership and spillover rates. (See, for example, KEMA, Inc. [2010].)

Evaluations can also include market actor survey data to estimate nonparticipant spillover and market effects. An important issue related to the quantification of nonparticipant spillover savings using only surveys of consumers is valuing the savings of measures installed outside the program. As previously noted, during telephone interviews consumers often cannot provide adequate equipment-specific data on new equipment installed either through or outside a program. Although they can usually report what type of equipment was installed, consumers typically cannot provide sufficient information about the quantity, size, efficiency, and/or operation of that equipment to enable a determination about its program eligibility.

One approach to estimating nonparticipant spillover and market effects via market actors is to ask market actors questions such as:

- What percentage of their sales meets or exceeds the program standards for each program measure category installed through the program(s)?

[43] There are studies that focus on examining how a change in the price of an energy-efficient product influences consumer purchases. Two approaches were used: (1) stated preference experiments that systematically ask potential consumers what they would choose from a set of options with different features and prices; and (2) revealed preference studies observe the actual choices consumers make from true choices available to them when making purchases. To obtain accurate revealed preference information, it is usually necessary to observe the items purchased. Consumers cannot reliably report the efficiency levels of recently purchased equipment. Direct observation can be accomplished via store intercepts for small items such as light bulbs, or via onsite visits for large items such as refrigerators. The remaining challenge for this method is the potential nonresponse bias; that is, potential differences between consumers who are willing to have their purchases observed and those who decline. An example of a study that focuses on how changes in price influence consumer purchases of energy efficient products is Cadmus (2012b). See the Appendix for additional information.

- What percentage of these sales did not receive an incentive?

The market actors should then be asked several questions about the program's impact on their decisions to recommend and/or install this efficient equipment outside the program.

4.3.4 Case Studies for Estimating Net Savings Using Survey Approaches

This section presents examples of estimating net savings with self-report surveys. Because self-report surveys are one of the most commonly used approaches, we provide four examples in this section. The first example demonstrates how the participant self-reports method is used to calculate free-ridership of residential and nonresidential programs in Illinois. The second example draws from work in California where self-report surveys are used to estimate free-ridership in nonresidential programs. A third example demonstrates how a sample set of survey questions were used in conjunction with a matrix to estimate free-ridership. The final example summarizes an approach used by the Energy Trust of Oregon (Castor 2012) that calculates low, mid, and high scenario NTG ratios to account for "Don't Know" responses to certain questions. This example addresses the best practice of conducting sensitivity analysis on the algorithm used to estimate NTG.

Example 1. Residential and Nonresidential Programs Free-Ridership Assessment

As part of a literature review for the Massachusetts Program Administrators to examine recent efforts to standardize measurement of net savings,[44] the evaluation team reviewed the recent efforts in Illinois to obtain consistent NTG methods. Below we summarize the background of this effort and the resulting recommended methods in both the residential and nonresidential program areas.

The Illinois (IL) Commerce Commission directed their evaluation teams to compile and formalize consistent NTG methods for use in IL EM&V work. The Commission's directives were twofold: (1) assess NTG methodologies and survey instruments that have been used to evaluate energy efficiency programs, and (2) compile the most justifiable and well-vetted methodologies in an attachment to the updated Illinois Technical Reference Manual (TRM) (Illinois Energy Efficiency Stakeholder Advisory Group 2016). The Commission noted that the IL NTG Methods should be flexible and adaptable to multiple program designs and budgets. It also noted the Methods should be tailored to appropriately assess the specifics of each of the Program Administrators' energy efficiency programs. The resulting statewide NTG methodology document covers the majority of residential and nonresidential programs offered in IL. If the NTG protocol is no longer appropriate, instructions are included for diverging from the IL NTG Methods.

Overview of Residential NTG Approaches—Illinois TRM

The Illinois TRM includes a residential cross-cutting NTG protocol as well as protocols for specific residential programs, including Appliance Recycling, Upstream Lighting, Prescriptive Rebate, Single Family Home Energy Audit, and Residential New Construction. The cross-cutting residential protocol formulates the core NTG as 1 – free-ridership (FR) + participant spillover

[44] Tetra Tech, NMR, and DNV GL, Net-to-Gross Methodology Research, prepared for the Massachusetts Program Administrators, March 24, 2017.

(PSO) and provides specific questions and scoring algorithms for measuring FR and PSO. It also provides specific questions and algorithms for measuring nonparticipant spillover (NPSO) from trade allies and customers, implying that they are to be included in the core NTG formula. This cross-cutting protocol provides detail on measuring PSO and NPSO, but it defers to the specific program protocols for measuring FR.

The specific protocols for programs include:

- The Appliance Recycling protocol includes basic and enhanced self-report approach (SRA) methods with specific questions and scoring algorithms for measuring FR including questions on how the appliance would have been disposed of in the absence of the program. The enhanced method may include additional research methods such as a retailer survey, appliance market assessment survey, or nonparticipant survey. The protocol does not provide specific guidance for when to use each SRA method nor for measuring SO, and thus the cross-cutting protocols may be assumed to prevail.
- The Residential Upstream Lighting protocol recommends using store intercept surveys for the customer SRA to measure PSO and NPSO. The protocol includes specific questions and scoring algorithms for measuring these NTG components. It includes specific questions for measuring FR and allows for partial FR. These include questions to assess program influence (captures the maximum level of program influence, reported by a survey respondent, of the residential lighting program on their decisions to purchase program bulbs on the day of the survey) and no-program questions (used to estimate how many program bulbs a survey respondent would have purchased in the absence of the residential lighting program); FR is calculated as the average of the responses to the two questions.
- The Prescriptive Rebate with No Audit protocol provides basic and enhanced methods with specific questions and scoring algorithms for measuring FR. Questions include program influence and no-program components[45] as well as consistency check questions on the program's influence to resolve possible conflicting responses. The basic method measures FR using a customer SRA. The enhanced method provides a protocol to triangulate and develop a weighted combination of FR estimates from two sources: the basic method and a trade ally survey. When multiple methods are used, evaluators may triangulate results by rating the analysis methodology and data collected using responses (rated on a scale of 0 to 10) to three questions: how likely is the approach to provide a more accurate estimate of FR, how valid is the data collected and the analysis performed, and how representative is the sample. The weight for each method is the average score for that method divided by the sum of the scores for all methods.
- The Single-Family Home Energy Audit protocol provides specific questions and scoring algorithms for measuring FR with different approaches for free/direct install versus rebated/discounted measures. The protocol measures FR using a customer SRA with questions on installation timing, quantity, and no-program scenario. Program influence

[45] Respondents are asked to report their likelihood (using a 0 to 10 scale where 0 is "not at all likely" and 10 is "extremely likely") to implement specified energy efficiency measures in the absence of the program. That likelihood score is then divided by 10 to produce the no-program score.

questions are excluded for free/direct install and included for rebated/discounted measures. It also includes consistency check questions on the program's influence for rebated/discounted measures to resolve possible conflicting responses.

- The Residential New Construction protocol recommends using builder surveys for the participant SRA to measure FR, PSO, and NPSO. The protocol includes specific questions and scoring algorithms for measuring these NTG components. The protocol measures FR using a participant SRA with questions on program influence installation timing, quantity, and no-program scenario. It also includes consistency check questions on the program's influence to resolve possible conflicting responses. PSO includes additional questions to help estimate the amount of savings using IL TRM protocols, such as quantity of appliances or location and amount of insulation. NPSO is based on surveys of two groups: dropout builders not participating in the past 12 months and true nonparticipating builders.

Overview of Nonresidential NTG Approaches—Illinois TRM

The IL TRM includes a core NTG protocol for nonresidential programs as well as protocols for specific programs, including the C&I New Construction, Small Business, and Study-based programs (e.g., programs that include an energy audit or assessment). There are core protocols for FR, PSO, and NPSO that provide specific questions and scoring algorithms associated with calculating FR and SO scores. That said, the core NTG ratio for an energy efficiency program is defined as 1 – FR even though they define PSO and NPSO.

The core FR protocol comprises three scores: Program Components FR Score, Program Influence FR Score, and No-Program FR Score, each ranging from 0 (no FR) to 1 (full FR). The three scores are combined to calculate the FR value. They are calculated as follows:

Program Components FR Score: Participants are asked to rate the importance of various factors on the decision to implement energy efficiency measures. The numeric scales range from 0 to 10, where 0 means "not at all important" and 10 means "extremely important." The factors included in the survey are program and non-program factors that could impact the participant decision-making process. The evaluator can calculate the score in one of two ways:

1. Equal to 1 - ([Maximum Program Factor Rating]/10).
2. Equal to 1 - ([Maximum Program Factor Rating]/([Maximum Program Factor Rating]+[Maximum Non-Program Factor Rating])).

Program Influence FR Score: Respondents are asked to allocate 100 points to the program and to non-program factors. The points the participants allocate to the program are the "Program Points." The "Program Influence FR Score" is calculated as 1 - (Program Points/100).

No-Program FR Score: Respondents are asked to report their likelihood (using a 0 to 10 scale where 0 is "not at all likely" and 10 is "extremely likely") to implement specified energy efficiency measures in the absence of the program. That likelihood score is then divided by 10 to produce this score.

The TRM states that consistency checks should be included in the survey questions to check the consistency of the FR responses. The protocol also provides guidance around vendor influence, including when and how to incorporate vendor responses into the FR calculation. The TRM outlines three scenarios to help decide when to utilize vendor responses, which is based on how involved the trade allies are in the program (i.e., integral in the delivery; part of a select, pre-approved network; implement projects and submit applications on behalf of the customer; sign agreements with the program administrator; or complete program-sponsored training). If vendor surveys are used, the TRM outlines questions that can be asked and based on the responses, when the results would be incorporated. Based on three scenarios, the evaluator decides if the vendor rating should be considered a program factor or non-program factor.[46]

- The Small Business protocol follows the core nonresidential FR protocol but includes a few exceptions primarily to reduce respondent burden.
- The C&I New Construction protocol follows the core nonresidential FR protocol but removes the timing aspect, as the program typically does not impact the acceleration of the construction.
- The Study-based protocol follows the core nonresidential FR protocol but includes additional questions about maintenance and performance of the measure.

The residential cross-cutting protocol states that FR questions should be asked near the beginning of the participant survey, before satisfaction questions. It also states that when estimating SO based on trade ally surveys, respondents should be allowed sufficient time to collect data to inform their responses and not rely on guesses.

The nonresidential core protocol does not provide direction on the timing of the FR survey. However, for SO, the protocol states the PSO module can be implemented as part of the NTG survey or separately, but timed to allow sufficient time—a minimum of three months—after program participation to allow for SO to occur.

Example 2. Nonresidential Programs Free-Ridership Assessment

The Large Nonresidential Free-Ridership Approach, developed by the Nonresidential Net-to-Gross Ratio Working Group for the Energy Division of the California Public Utilities Commission (2012), was developed to address the unique needs of large nonresidential customer projects developed through EE programs offered by the four California investor-owned utilities and other third parties. The Large Nonresidential Free-Ridership Approach is based on an approach that has been evolving for more than 15 years. As described in the framework, the method relies exclusively on the self-report approach to estimate project- and program-level NTG ratios, because the working group notes that other available methods and research designs are generally not feasible for large nonresidential customer programs. This methodology provides a standard framework, including decision rules, for integrating findings from quantitative and qualitative information in the systematic and consistent calculation of the NTG ratio.

The approach describes three levels of free-ridership analysis. The most detailed level of analysis, the Standard – Very Large Project NTG ratio, is applied to the largest and most

[46] Illinois Statewide Technical Reference Manual for Energy Efficiency, Version 5.0. Volume 4: Cross Cutting Measures and Attachments. February 11, 2016, Page 34.

complex projects (representing 10%–20% of the total projects) with the greatest expected levels of gross savings. The Standard NTG ratio, involving a somewhat less detailed level of analysis, is applied to projects with moderately high levels of gross savings. The Basic NTG ratio is applied to all remaining projects.

Five potential sources of free-ridership information are discussed in this study. Each level of analysis relies on information from one or more of these sources:

- **Program files**, which can include various pieces of information relevant to the analysis of free-ridership. Program files may include letters written by the utility's customer representatives that document what the consumer had planned to do in the absence of the rebate and explain the consumer's motivation for implementing the EE measure. It can also include information on the measure payback with and without the rebate.
- **Decision-maker surveys**, conducted with the person involved in the decision-making process that led to the implementation of measures under the program. This survey obtains highly structured responses concerning the probability that the consumer would have implemented the same measure in the absence of the program.
    - Participants are asked about the timing of their program awareness relative to their decision to purchase or implement the EE measure.
    - They are asked to rate the importance of the program versus non-program influences in their decision-making.
    - They are asked to rate the significance of various factors and events that may have led to their decision to implement the EE measure at the time that they did (for example, age or condition of the equipment, information from a facility audit, standard business practices, and experience with the program or measure).

  The survey also asks participants to describe what they would have done in the absence of the program, beginning with whether the implementation was an early replacement action. The decision-makers are asked to describe the equipment they would have installed in the absence of the program, including the efficiency levels and quantities. This information is used to adjust the gross engineering savings estimate for partial free-ridership.

  This survey contains a core set of questions for Basic NTG ratio sites, and several supplemental questions for both Standard and Standard – Very Large NTG ratio sites. For example, if Standard or Standard – Very Large respondents indicate that a financial calculation entered highly into their decision, they are asked additional questions about their *financial criteria* for investments and their rationale for the current project. These questions are intended to provide a deeper understanding of the decision-making process and the likely level of program influence versus these internal policies and procedures. Responses to these questions also serve as a basis for consistency checks to investigate conflicting answers about the relative importance of the program and other elements in influencing the decision. Standard – Very Large respondents may also receive additional detailed probing on various aspects of their installation decision based on industry- or technology-specific issues, as determined by review of other information sources. For

Standard – Very Large sites, the respondent data are used to construct an internally consistent "story" that supports the NTG ratio calculated, based on the overall feedback.

- **Vendor surveys** are completed for all Standard and Standard – Very Large participants who used vendors, as well as for Basic participants who indicate a high level of vendor influence in the decision to implement the EE measure. For participants who indicate the vendor was very influential in decision-making, the vendor survey results are incorporated directly into the NTG ratio scoring.
- **Utility and program staff interviews** for the Standard and Standard – Very Large NTG ratio analyses. Interviews with utility staff and program staff are also conducted to gather information on the historical background of the consumer's decision to install the efficient equipment, the role of the utility and program staff in this decision, and the names and contact information of vendors involved in the specification and installation of the equipment.
- **Other information** for Standard – Very Large Project NTG ratio sites includes secondary research of other pertinent data sources. For example, this could include a review of standard and best practices through industry associations, industry experts, and information from secondary sources (such as the U.S. Department of Energy's Industrial Technologies Program's Best Practices website). In addition, the Standard – Very Large NTG ratio analysis calls for interviews with other employees at the participant's firm, sometimes in other states, and equipment vendor experts from other states where the rebated equipment is installed (some without rebates) to provide further input on standard practice within each company.

Table 4 shows the data sources used in each of the three levels of free-ridership analysis. Although more than one level of analysis may share the same source, the amount of information used in the analysis may vary. For example, all three levels of analysis obtain core question data from the decision-maker survey.

Table 4. Information Sources for the Three Levels of NTG Ratio Analysis

	Program File	Decision-Maker Survey Core Question	Vendor Surveys	Decision-Maker Survey Supplemental Questions	Utility and Program Staff Interviews	Other Research Findings
Basic NTG ratio	√	√	√[a]		√[b]	
Standard NTG ratio	√	√	√[a]	√	√	
Standard NTG ratio—Very Large Projects	√	√	√[c]	√	√	√

[a] Performed only for sites that indicate a vendor influence score greater than maximum of the other program element scores.
[b] Performed only for sites that have a utility account representative.
[c] Performed only if significant vendor influence is reported or if secondary research indicates the installed measure may be becoming standard practice.

Example 3. Free-Ridership Assessment for an Equipment Rebate Program

This example shows how to calculate an NTG ratio and how to use a sample set of survey questions in conjunction with a matrix to estimate free-ridership (see Table 5). The example is from Chapter 5 of the Energy Efficiency Program Impact Evaluation Guide (SEE Action 2012b). In this case, the evaluators assign a free-ridership score based on a participant's response to six questions.

Table 5. Example Assignment of Free-Ridership Score Based on Participant Responses

Free-Ridership Score	Already Ordered or Installed	Would Have Installed Without Program	Same Efficiency	Would Have Installed All the Measures	Planning to Install Soon	Already in Budget
100%	Yes	Yes	—	—	—	—
0%	No	No	—	—	—	—
0%	No	Yes	No	—	—	—
50%	No	Yes	Yes	Yes	Yes	Yes
25%	No	Yes	Yes	Yes	No	Yes
25%	No	Yes	Yes	Yes	Yes	No
0%	No	Yes	Yes	Yes	No	No
25%	No	Yes	Yes	No	Yes	Yes
12.5%	No	Yes	Yes	No	No	Yes
12.5%	No	Yes	Yes	No	Yes	No
0%	No	Yes	Yes	No	No	No

*Source: SEE Action (2012b) based on example provided by Cadmus.

One issue with this method is the somewhat arbitrary nature of assigning free-ridership scores based on sets of question responses, as they depend on the judgment of the evaluator. Different researchers may assign different free-ridership scores to different sets of respondent answers. To address this, the literature recommends using sensitivity analyses around the free-ridership scores, based on the judgments of people familiar with the program.[47] An example of increasing the robustness of this method is found in an assessment of residential heating and cooling equipment for the Electric and Gas Program Administrators of Massachusetts.[48] Another useful exercise is to assess the reliability of the assignment of free-ridership scores by the evaluators. Inter-rater reliability scores[49] can be calculated to assess the reliability of these assignments. To the extent that evaluators assign the same free-ridership scores to the same set of response

[47] Issues may arise if these free-ridership scores are viewed as categories rather than as continuous variables. A 50% score may imply a higher level of free-ridership than does a 25% score, but it may not denote that the 50% score implies that free-ridership is, in fact, twice as high compared to respondents placed in 25% free-ridership score category. It is possible to perform arithmetic on these numbers and use the values to generate a mean value and even a variance, but this may not be appropriate. The lack of an accurate "distance" factor in these numbers makes the calculated variance hard to interpret. For variables that are meant to represent categories rather than continuous numeric values, frequencies are the more often used descriptive statistic.

[48] This work was conducted by a consortium of consultants under a prime contract led by Cadmus, supported by Navigant, and Opinion Dynamics Corporation (cited as Cadmus; Navigant Consulting; Opinion Dynamics Corporation (2012).

[49] *Inter-rater reliability*, *inter-observer reliability*, and *inter-judge agreement* are some terms that have been used in the literature to designate a wide variety of concepts. All these terms, however, refer to the extent of agreement among raters, judges, and observers (Gwet 2010, 2012).

patterns, then reliability will be increased. Other approaches use upper and lower bounds on free-ridership developed directly from survey respondents.[50]

Example 4. Commercial, Industrial, and Residential Scenario Analysis

The Energy Trust of Oregon uses an approach (Castor 2012) to calculate low, mid, and high scenario NTG ratios to account for the "Don't Know" responses to certain questions. The report appendix describes this approach. The project's free-ridership score is composed of two elements: a project change score and an influence score.

The project change score is based on the respondent's answer to the question, "Which of the following statements describe the actions you would have taken if Energy Trust incentives and information were not available"? Possible answer choices are assigned a number between 0 and 0.5, with 0 indicating no free-ridership and 0.5 indicating that the participant was a full free-rider. Because a respondent can select multiple responses to the question, the answer choice with the lowest score is selected. If the respondent selects "Don't Know," two scores are created to account for the range of possible answers (0 and 0.5).

For commercial projects, respondents are asked this follow-up question when they report they would not have done anything differently in the absence of the program: "If your firm had not received the incentive, would it have made available the funds needed to cover the entire cost of the project"? If the respondents select "Yes," their project change score is 0.5. If the respondents select "No," their project change score is 0. However, if the respondents select "Don't Know," they are given two scores for project change, as previously described.

The influence score is based on respondents' answers to questions about the influence of Energy Trust incentives, program representatives, contractor/salesperson, studies, and other program elements. The answer choices are given a value between 0 (element's influence was a 5, extremely influential) and 0.5 (element's influence was a 1, not at all influential). The score for the most influential element is taken as the influence score. If respondents answer "Don't Know" for all elements, they are given two influence scores to account for the range of possible answers (0 and 0.5).

[50] Violette et al. (2005) discuss approaches used in the net savings and attribution assessment for a large-scale C&I retrofit program. Free-ridership was assessed using a series of survey questions asked of various actors, including participating end-use consumers and vendors/contractors/consultants. Free-ridership was asked in direct free-ridership questions and supporting, or influencing, questions. Participating owners and energy service companies/contractors in a large-scale C&I retrofit program were each asked for direct estimates of: (1) the "proportion" of the savings or measures that would have been installed without the program; and (2) the "likelihood" that the measures would have been installed without the program. A three-step approach was used. Step 1 focused on whether the respondent believed that free-ridership existed at all; if the respondent believed it existed in this project, Step 2 established bounds on the free-ridership effect, that is, what was the smallest value that seemed reasonable and what might have been the highest reasonable free-ridership value. Step 3 used questions to obtain where within this range the free-ridership value was likely to fall. Appendices to Violette et al. (2005) discuss alternative approaches. This program had some unique characteristics that made this approach more tractable. It involved large-scale C&I projects and the survey respondents were provided with summaries of the technologies and measures installed. Other efforts that used similar approaches include Violette, Ozog and Cooney (2003) for addressing net savings from regional and market transformation programs in the Pacific Northwest, and Navigant (2013b) which assesses the net impacts of U.S. DOE's Wind Powering America Initiative.

To generate the free-ridership score for each project, the project change and influence scores are added. For respondents who do not provide "Don't Know" answers, this score will be a single number between 0 (no free-ridership) and 1 (full free-ridership). For those who gave a "Don't Know" answer to one of the questions, there are two free-ridership scores—one high and one low. For those who answered "Don't Know" to *both* the project change and influence questions, no score is calculated.

Free-ridership scores are averaged for all respondents in each program/measure group and the result is shown as a percentage rather than a decimal (see Table 6 for pros and cons of survey-based approaches).

- "Low Scenario" is the average of the free-ridership scores where the low score is used for those who answered "Don't Know" to a question.
- "High Scenario" is the average where the high score is used for those who answered "Don't know" to a question.
- "Mid Scenario" is the average of the Low and High Scenarios. In the case of C&I projects, individual scores are weighted by their share in the electricity or gas savings of all respondents of their group before the scores are averaged for scenarios.

Table 6. Survey-Based Approaches—Summary View of Pros and Cons

Pros	Can provide useful information to support process and impact evaluations (for example, source of awareness, satisfaction, and demographics) Flexible approach that allows the evaluator to tailor questions to the program design or implementation methods Participant self-reports can yield estimates of free-ridership and spillover without the need for a nonparticipant control group Nonparticipant and market actor interviews can be used to triangulate participant self-report responses and calculate nonparticipant spillover or market effects.
Cons	Potential biases related to respondents' giving "socially desirable" answers Consumers' inability to know what they would have done in a hypothetical alternative situation, especially in current program designs that use multiple methods to influence behavior The tendency of participants to rationalize past choices Potential arbitrariness of scoring methods based on evaluator judgment that translate responses into free-rider estimates Participants may fail to recognize the influence the program may have had on other parties who influenced their decisions (for example, program may have influenced contractor practices, which in turn impacted the participant) Participant surveys capture only a subset of market effects Amount of time and cost to identify a group of nonparticipants who have installed energy-efficient equipment on their own.

4.4 Market Sales Data Analyses (Cross-Sectional Studies)

A market sales data method can capture the total net effect of the program, including both free-ridership and participant and nonparticipant "like" spillover. As described in a residential free-ridership and spillover methodology study prepared for the Massachusetts Program Administrators (NMR Group, Inc. and Tetra Tech 2011), the total net effects of a program can be estimated via an analysis of market sales data.

The most common approach is a cross-sectional comparison area method in which post-program data are compared with data from a non-program comparison area (or multiple comparison areas) for the same point in time. Thus, evaluators can make a comparison between the change in the program area from the pre-program period to the post-program period *and* the change in the non-program area over the same period.

The NMR Group, Inc. and Tetra Tech (2011) study lists three important factors to consider when deciding if an approach is appropriate for a program:

- **Does an appropriate comparison area exist?** Comparison area(s) must represent a credible baseline for the area of interest. This may entail using a set of systematic adjustments to control for differences in total size of, or demographics for, the areas. As EE programs become more prevalent, finding comparison areas that do not have similar program activities is becoming more difficult.
- **Are the market data available and complete?** Market data analysis requires comprehensive market data for the area of interest and an appropriate comparison area or areas. The complication here is that comprehensive sales/shipment tracking systems have not been available for most markets. Absent comprehensive sales data, a general picture of market coverage can be obtained by conducting surveys or in-depth interviews. These are typically conducted with vendors and contractors about sales volumes and efficient equipment sales shares for conditions with and without the program, or for in-territory and comparison area sales. In some cases, the self-reported purchases of participating end users can provide market data if the sample is sufficiently large and representative of the market. Also, it can be expensive to gather the market sales and shipment data, and even a diligent data collection effort may leave gaps in the data.
- **What are the features of the program?** Market data analysis is usually appropriate for programs that promote large numbers of homogenous measures and that have substantial influence upstream to the end user.

As an example of this approach, Cadmus et al. (2012) tracked ENERGY STAR® appliances, lighting, and home electronics product sales in New York and then compared those sales to sales of the same products in Washington, D.C., Houston, Texas, and Ohio. All these baseline areas were without significant utility efforts to promote ENERGY STAR products. The market data were used to estimate both the market share and the energy savings attributable to the New York Energy $mart Products Initiative Program administered by the New York State Energy Research and Development Authority.[51]

[51] Scott Dimetrosky indicated that this study developed savings from product sales and installations. These savings were derived by first estimating the market share for ENERGY STAR products through estimates of total market size and sales of ENERGY STAR products. Next, portions of the market share were allocated to exogenous, non-New York Energy $mart Products Initiative Program (NYE$P) effects, including the impact of the national U.S.

Another example of a market sales approach entails interviewing or surveying a panel of trade allies who are either program participants or nonparticipants. This could include contractors, retailers, builders, and installers. These trade allies are offered monetary compensation for information about projects or sales completed within a specified time period (see Table 8for pros and cons of this approach). The types of information requested can include manufacturer, efficiency levels, size, price, installation date, installation ZIP code, types of incentives received, and an assessment of the program's impact on incented and non-incented efficiency actions. With annual updates, this method could provide context for tracking longer term ongoing program impacts or market effects. This method could also work in tandem with other approaches for estimating net savings and provide a market context for estimates that may otherwise focus only on short-term impacts.

Another more detailed example of a recent market sales data analysis using in-store visits and web scraping is shown below.

4.4.1 Case Study for Market Sales Data Analysis

Example 1: Massachusetts RLPNC 16-6: Lighting Shelf Stocking

On behalf of the Massachusetts ENERGY STAR Lighting program administrators and Lockheed Martin, NMR Group, Inc. conducted a shelf-stocking and price survey to evaluate the impact of the Mass Save® residential lighting program on consumer retail lighting in Massachusetts (NMR 2017).

The study took advantage of two separate but complementary data collection methods: 1) site visits to 100 stores in Massachusetts and 30 stores in New York in 2016 to inventory light emitting diode, compact fluorescent lamp (CFL), incandescent, and halogen lamp packages, and 2) web scraping, which provided time series data on lighting cost and availability through the collection of data from retailer web pages. The authors noted that while shelf-stocking studies provide a useful look at lighting cost and availability at a discrete point in time, web scraping adds time series data on lighting cost and availability in the marketplace over time. Because the study used two methods, the authors could compare the data collected through both methods and learn how online and in-store prices and availability differed.

The authors noted that both methods offered distinct advantages. Physically visiting stores is the only way to learn information about how products are displayed in the store, the amount of shelf space given to different products types, and what indirect and direct signals stores are providing to customers about the value and desirability of the products. Web scraping offered a number of other advantages:

- Eliminated the financial and time cost of travel, training, and obtaining permission to visit retailers

Environmental Protection Agency/U.S. Department of Energy ENERGY STAR Program, naturally occurring adoption (including the impact of higher energy prices and interest generated by programs in neighboring states), and the impacts of other New York State Energy Research and Development Authority residential programs. The remaining market share, after netting out these other effects, was considered attributable to the NYE$P.

- Not particularly difficult or expensive to set up as the standard methods use free open-source software that is widely used and well documented
- Scraping can be easily automated to run on a regular schedule to create rich time-series datasets.

On the other hand, web scraping had some inherent caveats and limitations:

- The information available is only as good as the websites' administrators make it
- Websites tend to change frequently, which requires updates to the code
- Markdown and rebate information is included inconsistently
- There is not a way to verify how online products offerings, prices, and stock data correspond to what is actually in a store
- The amount of data generated results in the need to filter and clean the data to generate useful insights.

The overall conclusion of the study was that incorporating both data streams offered a richer picture of the lighting market during the study period.

The pros and cons of market sales data analyses are listed in Table 7.

Table 7. Market Sales Data Analyses—Summary View of Pros and Cons

Pros	Can estimate the total net effect of a program Uses information on actual consumer behavior Addresses trends in an entire market Most appropriate for programs that promote large numbers of homogeneous measures and have substantial influence upstream.
Cons	There may be a low availability and quality of sales and shipment data in the area of interest and in an appropriate comparison area(s) Data may be expensive to acquire and/or may have gaps that can be misleading May be difficult to determine the appropriateness of a comparison area.

4.5 Structured Expert Judgment Approaches

Structured expert judgment approaches involve assembling a panel of experts who have a good working knowledge of the technology, infrastructure systems, markets, and political environments. This approach is one alternative for addressing market effects in different end-use markets. These experts are asked to estimate baseline market share for a measure or behavior. In some cases, they are also asked to forecast market share with and without the program in place. Structured expert judgment processes use a variety of specific techniques to ensure that the panel of experts specify and take into account key known facts about the program, the technologies supported, and the development of other influences over time (Tetra Tech et al. 2011).

The Delphi process is the most widely known technique (NMR Group, Inc. and Research Into Action 2010). Each panelist is asked to make a judgment on the topic—based on the provided information and on his or her experience—and submit the information to the evaluators. The evaluators compile the information from the panelists and return it to the panelists for another review. The panelists are asked whether they stand by their original judgments or whether the assessments of their peers have caused them to alter their judgments. At least two rounds of judgment are required for a Delphi panel, although more rounds can be used.

Some advantages of the structured expert judgment approach are:

- The estimate is based on feedback from a group of experts, which can be particularly useful for programs with complex end uses.
- It is a useful tool for consolidating results from multiple methods to develop a consensus estimate (see example 2 below).

As with other approaches (such as market sales data analysis), the structured expert judgment method relies on high-quality data to inform the panel, so sparse data can result in inaccurate estimates of net savings (NMR Group, Inc. and Research Into Action 2010).

Two examples of using the structured expert judgment approach to estimate net savings are presented here. The first example describes how Delphi panels were used to estimate net savings for a residential new construction program in California. The second example describes the development a final estimate using a Delphi panel's review of estimates.[52]

4.5.1 Case Studies for the Structured Expert Judgment Approach

Example 1: Residential New Construction Delphi Panel

In a study prepared for the California Public Utilities Commission Energy Division, evaluators used two Delphi panels of Title 24 consultants and building industry experts to convert the gross savings estimates. The panel converted estimates from investor-owned utilities (IOU) programs targeting the residential new construction sector to net savings estimates (Hoefgen et al. 2011).

The panelists received detailed data pertaining to code compliance, compliance margins, and estimates of annual gross energy savings in non-program homes at the state level and by climate region. After reviewing these data, panelists were asked to:

- Estimate the proportion of the electricity and natural gas savings attributable to the IOU programs targeting the residential new construction sector and other factors (non-IOU residential new construction programs, the economy/housing market, energy prices, and climate change).
- Estimate the percentage of net savings in non-program homes attributable to different IOU program elements (builder trainings, incentives, and design assistance).

[52] An application of the Delphi technique as applied outside of EE may be informative. Navigant (2013b) conducted an evaluation of the Wind Power America program. The goal was to assess the impacts attributable to the program. The unique aspect of this Delphi exercise was the use of range estimates; that is, experts were asked about lower and upper bounds to the effects as well as a best estimate. This approach allowed the experts to provide their own insights into the uncertainty of the estimates. Gauging uncertainty and then using that in probabilistic and scenario analyses are consistent with other utility resource planning activities. Adapting these methods to EE resource assessment may increase the usefulness of the information.

- Assess the extent to which the market effects were likely to persist in the absence or reduction of the IOU programs.
- Estimate the percentage of homes that would have been below code in the absence of the IOUs' programs and other factors, and estimate the compliance margin of the below-code homes in the absence of each factor.

Each panelist completed two rounds of detailed surveys. In the second round, they were provided a comparison with other panelists' responses and logic and allowed to change their answers. The evaluation team analyzed the Title 24 consultant responses (both weighted and unweighted) using the building industry experts' responses as a qualitative check. The Delphi panel provided estimates on gross electricity and gross natural gas savings from above-code homes. Both panels identified the various elements of training (builders, subcontractors, and Title 24 and code officials) as the most important elements of the IOUs' programs.

Example 2: Lighting Program Delphi Panel

Another way to use a Delphi panel is to have the panel review estimates derived through other methods to develop a final estimate. As part of the evaluation of the Massachusetts ENERGY STAR Lighting Program (KEMA 2010), evaluators used a Delphi panel of lighting and EE experts across the United States and Canada. The panelists were asked to integrate results from five methodologies that yielded NTG estimates (conjoint analysis, multistate modeling, revealed preference study, supplier interviews, and a willingness-to-pay study). Evaluators then used the Delphi panel's review in developing recommendations for the final NTG estimate.

See Table 8 for pros and cons of the structured expert judgment approach.

Table 8. Structured Expert Judgment Approaches—Summary View of Pros and Cons

Pros	The resulting estimate is the independent, professional judgment of a group of technology and/or market experts It is a useful approach for programs with diverse and complex end uses or practices Is a useful tool for consolidating results from multiple methods to develop a consensus estimate Panel members can provide levels of confidence and procedures using appropriate elicitation methods.
Cons	The approach relies on high-quality data to inform the panel, leading to reasonable estimates of net savings Sampling-based calculations of confidence and precision are not available The approach is judgmental/subjective.

4.6 Deemed or Stipulated Net-to-Gross Ratios

Deemed or stipulated NTG ratios are predetermined values and do not rely on a calculation-based approach. Deemed values are often based on previous NTG research that was conducted using at least one of the other methods described in this chapter.

NTG ratios are often stipulated when the expense of conducting NTG ratio analyses cannot be justified or when the uncertainty of the potential results is too great to warrant a study. A recent review of 42 jurisdictions in the United States and Canada (which represented nearly all jurisdictions with ratepayer-funded EE programs) found that only 14% use a deemed approach to NTG for C&I programs compared to 50% of the jurisdictions using an active research approach to developing estimates of net savings factors (Navigant 2013a).[53]

Deemed or stipulated NTG ratios are typically either set by a regulatory agency or negotiated between regulators and program administrators. These ratios may be determined at the portfolio level (for example, Michigan and Arkansas)[54] or on a measure-by-measure basis (for example, California and Vermont).[55] Typically, evaluators base the ratios on NTG studies from past evaluations and/or reviews of other similar programs in which a NTG ratio was estimated. For example, it is not unusual in a multiyear portfolio cycle to estimate a NTG ratio for an initial year (or possibly every other year), with deemed values used in the subsequent or intervening years. This multiyear estimation of NTG ratios is a compromise between performing net savings estimation studies every year and the use of deemed values based on that research for a selected time period. As an example, Massachusetts moved to this approach.[56]

In other cases, evaluators use historical data or other information from a wide range of sources to develop a "weight of evidence" conclusion about the program's influence (SEE Action 2012b). As discussed earlier, one common approach for developing a stipulated value is to use a panel of experts who have the relevant experience to make that judgment (Delphi panel).

Although using deemed or stipulated values is a relatively simple and low-cost approach, there are several disadvantages. NTG values are variable across time and space, and strongly linked to program design and implementation. This makes deemed values or assumptions potentially unreliable when transferred from a program in one jurisdiction to a similar program in another jurisdiction.[57] NTG values based on primary research efforts can produce estimates that are based on program-specific information (NMR Group, Inc. and Research Into Action 2010). As a

[53] Approximately one third of the jurisdictions did not adjust gross savings for either free-ridership or spillover; however, many of those states conducted some NTG research to inform future program design. This reflects policy decisions in each state. Several states that did not adjust gross savings for net savings factors at the time of this study have changed or are contemplating changing to approaches that do estimate net savings. Pennsylvania and Maryland fall into this category. In Pennsylvania, Act 129 program savings targets are based on gross savings, but utilities participating in Act 129 programs are required to report gross and net savings.

[54] Arkansas: NTG deemed at 0.8, www.apscservices.info/pdf/07/07-085-tf_286_44.pdf; Michigan: NTG is deemed at 0.9 for all programs except pilot, education, and low-income programs, which are deemed at 1.0. http://efile.mpsc.state.mi.us/efile/docs/17138/0009.pdf. Note that most low-income programs are not subject to NTG analysis (that is, are deemed at 1.0).

[55] California Database for Energy Efficient Resources (DEER): http://www.deeresources.com/; Vermont, see: www.efficiencyvermont.com/docs/about_efficiency_vermont/annual_reports/2011_Gross_to_Net_Report_EfficiencyVermont.pdf.

[56] Massachusetts has been conducting extensive NTG research, but has moved to deemed/stipulated values for their 3-year plan. Any NTG variances from the stipulated values have no effect on current cost recovery or incentive payments. Yet the extensive program- and measure-level NTG research continues where appropriate, and the state is benefiting from improved program designs without major controversy involving cost recovery and incentives for current programs.

[57] Another issue raised by a reviewer was that the use of deemed NTG values can remove the incentive for the program administrator to reduce free-ridership and maximize spillover and market effects to yield greater net savings values.

result, these values provide useful information for the design and implementation of programs[58] and may mitigate the risk to ratepayers from utilities receiving performance incentive payments on savings not actually attributable to the program (as well as the risk to ratepayers of making performance incentive payments that are too large). NTG values are also critical from a resource planning perspective and having better data on the actual energy savings achieved from energy efficiency programs can help the planning process (Navigant 2013a). Deemed or stipulated NTG values do not provide these benefits.

The following example illustrates how one agency uses deemed savings for program planning.

4.6.1 Case Study for Using Deemed Savings

Example 1: California Public Utilities Commission Database for Energy Efficient Resources

The California Public Utilities Commission uses deemed savings (listed in its Database for Energy Efficient Resources) for planning purposes and interim savings estimates for its programs. These deemed savings are updated based on results of NTG studies. NTG savings values are presented for kilowatt-hours and kilowatts.

See Table 9 for pros and cons of a deemed savings approach.

Table 9. Deemed or Stipulated Approaches—Summary View of Pros and Cons

Pros	This approach can reduce contentious after-implementation adjustments to estimated program savings because agreed-on net savings factors are developed in advance of program implementation.
Cons	An incorrect estimate can be deemed It is not based on current program-specific information The evaluator cannot assign sample-based statistical precision to the estimate Developing deemed savings net values at the measure and technology levels can be time consuming and expensive The process for developing deemed net savings can be contentious.

4.7 Historical Tracing (or Case Study) Method

This method involves reconstructing the events (such as the launch of a product or the passage of legislation) that led to the outcome of interest. An example of this is developing a "weight of evidence" conclusion about the specific influence a program had on the outcome.

Historical tracing relies on logical devices typically found in historical studies, journalism, and legal arguments (Rosenberg and Hoefgen 2009). These include:

- Compiling, comparing, and weighing the merits of narratives of the same set of events provided by individuals who have different points of view and interests in the outcome

[58] For example, free-ridership can inform decisions to discontinue incenting certain measures, increase incentive amounts, or increase the efficiency level being incented.

- Compiling detailed chronological narratives of the events in question to validate hypotheses regarding patterns of influence
- Positing a number of alternative causal hypotheses and examining their consistency with the narrative fact pattern
- Assessing the consistency of the observed fact pattern with linkages predicted by the program logic model
- Using information from a wide range of sources (including public and private documents, personal interviews, and surveys) to inform historical tracing analyses.

The historical tracing method traces chronologically a series of interrelated events either going forward from the research point of interest to downstream outcomes, or working backward from an outcome along a path that is expected to lead to precursor events. If all likely paths are followed, forward tracing can capture a relatively comprehensive view of project or program effects. Because the path leads from a program event, the connection to the event is assured. Backward tracing usually focuses on a single outcome of importance and follows the trail back through developments that seem to have been critical to reaching the identified outcome. These developments may or may not link back to the research program of interest (see Ruegg and Jordan 2007).

Weiss (1997) suggests historical tracing is similar to theory-driven evaluation and can be viewed as an alternative to classical experimental design. This approach suggests that if the predicted steps between an activity and an outcome can be confirmed in implementation, this matching of the theory to the observed outcomes will lend a strong argument for causality. In other words, if the evaluation can show a series of microsteps that lead from inputs to outcomes, causal attribution, for all practical purposes, is supported by this approach.

Scriven (2009) argues that some researchers have been entranced by the paragon of experimental design—the RCT—and have generalized this into a virtual standard for good causal investigation. This view can be contrasted to the way that “epidemiology, engineering, geology, field biology, and many other sciences establish causal conclusions to the highest standards of scientific (and legal) credibility” (p. 151).

This method is best suited to an attribution analysis of major events, such as adoption of new building codes or policies. It is not typically applicable to EE programs. However, various elements of this approach may be used in the analysis of very large custom projects that essentially require case study approaches.

Because this method draws from multiple information sources, it is difficult or impossible to determine the magnitude of the effects, so the evaluator cannot assign statistical precision to the estimate (NMR Group, Inc. and Research Into Action 2010). However, as part of making a persuasive case for attribution and providing evidence supporting a statistically derived net savings estimate, this method can be very important. Statistics alone often do not constitute a complete attribution assessment. They often require context using supporting logic to enhance the validity of the statistical estimates, as illustrated in the following example.

4.7.1 Case Study for Using the Historical Tracing Method

Example 1. Historical Tracing for a Residential New Construction Program

Keneipp et al. (2011) used historical tracing in conjunction with Delphi panels to develop energy savings for new homes (see Table 10 for pros and cons of this approach). This study used historical tracing spanning 14 years of regulatory documents to create timelines of the residential new construction program presence and activities for Arizona Public Service Company. The evaluators used these data to create an influence diagram of market influences on specific building practices. This information was then shared with two in-person Delphi panels of market experts who estimated the percentage of homes built in 2010 using specific building practices. These Delphi panels also developed the counterfactual scenarios used to show the net impact of the residential program on the percentage of homes that were built to standards, but would not have met these standards in the absence of the program. The Delphi outputs were then used to develop inputs for an engineering simulation model to calculate energy savings per home. This example illustrates how historical tracing can be used in combination with other methods to develop actual quantitative net savings estimates from an EE program.

Table 10. Historical Tracing (or Case Study) Method—Summary View of Pros and Cons

Pros	Draws from multiple information sources Can be used at a market level for upstream EE programs Can be useful for making a persuasive case for attribution and provide evidence to support a statistically derived net savings estimate.
Cons	It can be difficult to translate the influence factors into estimates of impacts without additional modeling The evaluator cannot calculate sample-based statistical confidence and precision levels for the estimate.

4.8 Common Practice Baseline Approaches

The common practice baseline approach[59] is also is receiving attention as a method for estimating net savings. SEE Action (2012b) has defined the common practice baseline as follows:

> *Common practice baselines are estimates of what a typical consumer would have done at the time of the project implementation. Essentially, what is "commonly done" becomes the basis for baseline energy consumption (SEE Action, 2012b, p. 7-2).*[60,61]

[59] The Common Practice Baseline section gave rise to several comments. Some reviewers did not see this method as parallel to the other methods presented in this chapter, as it focuses on *ex ante* values of the mean of market behavior and does not look at *ex post* information on actions or program participants. In this context, this approach was viewed as more of an *ex ante* deemed net savings approach (see Section 4.7 on deemed NTG values). After considering these comments, the Common Practice Baseline approach was viewed as warranting a separate section due, in part, to the recent attention given this approach to net savings.

[60] SEE Action (2012b) illustrates this "commonly done" baseline using an appliance example. "For example, if the program involves incenting consumers to buy high-efficiency refrigerators that use 20% less energy than the minimum requirements for ENERGY STAR® refrigerators, the common practice baseline would be refrigerators

This baseline includes a "consideration of what typically would have been done in the absence of the efficiency action" (SEE Action 2012b). This approach is under development in several jurisdictions and will certainly evolve in its application. In general, it is based on using available information to develop an *ex ante* estimate of net savings, with limited adjustments based on *ex post* data and analysis. This approach has many appealing qualities, but the tradeoffs need to be clarified, both in terms of potential biases and the real costs associated with this approach.

The common practice baseline method is relatively new in the broader evaluation literature and its application has been somewhat limited; however, the Northwest Power and Conservation Council (NW Council) in the Pacific Northwest has applied a variant of this method for a number of years in estimating *ex ante* net savings.[62] The NW Council continues to evolve this approach with new protocols developed by the Regional Technical Forum (RTF 2012).[63] Ridge et al. (2013) indicate that, in addition to the NW Council, three other jurisdictions are working with variants of the common practice baseline approach: Indiana, Delaware, and Wisconsin (Focus on Energy). In general, these jurisdictions have evaluation guidelines or regulatory framework that allows for the use of common practice baseline variants under certain circumstances, but they also allow for and use survey-based approaches and RCT or quasi-experimental design approaches to estimated net savings for many programs.

4.8.1 Common Practice Baselines—Discussion

As with other net savings approaches, the common practice baseline approach is designed to assess the savings attributable to EE program activities. One advantage claimed for the common baseline approach is that it avoids double counting of free-riders. The concern is that the two-step approach—where (1) gross savings is estimated *ex post* using a baseline that may be similar to "common practice"; and (2) an NTG ratio is applied to the *ex post* gross savings—can double count at least some free-riders (Ridge et al. 2013; Hall et al. 2013). The argument is that the *ex*

---

that consumers typically buy. This might be non-ENERGY STAR refrigerators, or ENERGY STAR refrigerators, or, on average, something in between."

[61] SEE Action (2012b) defines common practice baselines in its glossary as "The predominant technology(ies) implemented or practice(s) undertaken in a particular region or sector." (p. A-4).

[62] Tom Eckman of NW Council indicated that this general approach has been applied in setting deemed savings since the 1980s, and it was designed to fit with the NW Council integrated planning process; that is, it is meant to provide an estimate of the increment of savings beyond what system planners assume for naturally (or currently) occurring efficiency in their demand models. Additional information can be found at the RTF website of the NW Council and in RTF (2012) as well as in the roadmap for the assessment of EE measures (RTF, 2015).

[63] Some reviewers indicated that this double counting problem may be the result of inconsistent program rules as set out by the program administrators and regulators, not an estimation issue. Further, a number of reviewers indicated that rather than over-estimating free-riders, this approach underestimates free-riders due to selection bias (discussed in the main body text below). The RTF guidelines (dated August 15, 2012) sets out the current practice baseline approach most directly in its definition of savings: "Savings is defined as the difference in energy use between the baseline (see section 3.2) and post (after measure delivery) periods, which is caused by the delivery of a measure. The terms "net" or "gross" are intentionally not used to modify the term "savings," as they may conflict with the definition of "baseline," provided in section 3.2. The current practice baseline defines directly the conditions that would prevail in the absence of the program (the counterfactual), as dictated by codes and standards or the current practices of the market. The most important conflict would arise if savings were estimated against a current practice baseline and then those savings were further adjusted by a net-to-gross ratio, where the net-to-gross ratio was the probability that the measure would have been delivered in the absence of program influence." Note that the RTF uses the term *current baseline* rather than *common practice baseline* used elsewhere.

ante estimate of gross savings may be close to net savings without any adjustment for NTG factors such as free-ridership, spillover, and market effects. This view assumes that some of these NTG factors are already accounted for by the process used to produce the *ex ante* gross savings estimates. This emphasizes the need to: (1) understand the derivation of gross estimates as part of the EE evaluation process, and (2) to explicitly set out the assumed counterfactual scenario in both the gross savings and net savings methods used.[64] Taking these two steps avoids the double counting that results in higher-than-appropriate free-ridership estimates.[65]

Massachusetts has recently adopted a gross baseline framework (DNV GL and ERS 2017) designed to be consistent with the ISO-NE Forward Capacity Market requirement, where baseline is based on the more stringent of existing code/standard and "Industry Standard Practice." The transition plan to implement the new framework includes revising the net-to-gross survey process to ensure that net savings is neither over- nor under-estimated as a result of the gross baseline revision.

Examples from guidelines on common practice baselines include:

- **NW Council's guidelines savings estimation methods:** The NW Council through its Regional Technical Forum (RTF) has the longest history in using a common practice baseline approach. Termed "current practice baseline" by the RTF, this baseline defines directly the conditions that would prevail in the absence of the program (the counterfactual scenario), as dictated by codes and standards or the "current practices of the market." (RTF 2015, p. 3) with current practice defined as the "typical choices of eligible end users, as dictated by codes and standards and the current practices of the market." The RTF estimates this baseline based on recent choices of eligible end users in purchasing new equipment and services. These choices may be inferred from data on shipments, purchases (equipment or services) or selected design /construction features. For example, the baseline for more efficient televisions is the average efficiency of recent television shipments. These baselines along with the measure unit energy savings are subject to a sunset date. The sunset date is "shortened as needed to reliably estimate savings for a measure whose baseline is rapidly changing." (p.10). The RTF sets out indicators used to determine if current practice is the appropriate baseline. However, "as a general rule, the RTF will use a baseline that is characterized by current market practice or the minimum requirements of applicable codes or standards, whichever is more efficient." (p.10).

[64] It is important to remember that both gross savings and net savings are difference estimates and both need a baseline for estimation (see NEEP, 2016).

[65] Some reviewers indicated that this double counting problem may be the result of inconsistent program rules as set out by the program administrators and regulators, and is not an estimation issue. If this is the case, evaluators still must decide whether the *ex ante* savings are net, gross, or somewhere between, because the *ex post* estimates must be used in an internally consistent way to adjust the claimed *ex ante* savings. Further, a number of reviewers indicated that rather than overestimating free-riders, this approach is likely to underestimates free-riders because of selection bias (discussed in this section).

Indiana and Delaware evaluation frameworks: The evaluation guidelines developed in two state-wide frameworks[66] list the use of the standard market practice as approaches that can be used in the estimation of net savings in utility evaluation of EE programs. Indiana indicates that this approach is a way to set energy impact analysis baselines so that the baseline already incorporates the influence of free-riders. In this approach, a free-rider assessment is not needed because the market is already using a standard market practice baseline without the program's direct influence. This baseline is typically set at the mean of the level of EE being installed across the market being targeted by the program (TecMarket Works et al. 2012, p. 55). An update to the Delaware State-Wide Evaluation Framework (Optimal Energy, 2015) also listed standard market practice baselines as a candidate approaches for use in estimating net savings. "Because free-riders are expected to take part in Delaware programs, a Net-to-Gross analysis will be completed for all programs in which free-riders are expected, unless the evaluation approaches use experimental or quasi-experimental designs or set energy impact baseline conditions at standard market practice levels that lead directly to the estimation of net savings." In addition to the evaluation guidelines discussed above, the Wisconsin Focus on Energy (FOE) used a common practice baseline method for a residential program in recent evaluation work (FOE 2017) and sets out processes for use of this method in future evaluation work. The case for the use of a common practice baseline approach appears to stem largely from two issues:

1. The definition of gross savings may include factors that are more appropriately viewed as components of net savings, and additional adjustments are not needed to these original estimates. This is essentially an *ex ante* estimate of net savings using current practice as the baseline with net savings estimated as the reduction in energy use resulting from the change to more efficient technologies.[67,68]
2. Program evaluations that report net savings may do so inconsistently. Unfortunately, the components of the net savings calculation differ between jurisdictions, and are often based on what the jurisdiction's stakeholders view as appropriate and measurable (see

---

[66] These two state-wide frameworks provide guidance on evaluation methods for utility EE evaluations and include the use of common or standard practice baselines as candidate methods; however, it is not clear how often a common or standard practice baseline method has been selected for use by utilities in these states. An evaluation report addressing Indianapolis Power and Light's EE programs (2015) did not use CPB methods and instead used survey methods for estimating net savings in C&I and residential programs.

[67] Tom Eckman of the NW Council expands on this point, stating that, "What is occurring prior to program launch is a better measure of what would have occurred absent the program (that is, the counterfactual scenario) than a determination made after the program has influenced the market." Essentially, the NW Council performed an *ex ante* net analysis when they developed deemed savings estimates that are by design viewed as net savings. For the NW Council's purposes, this is viewed as being as accurate as performing complex studies after the program has been implemented. More information on the NW Council approach can be found in RTF (2012) and at the RTF website http://rtf.nwcouncil.org/.

[68] The common practice approach as applied by the NW Council works best when the forecasts are made at the measure level. Covering all the measures that combine to make a program can be time consuming and expensive to update. Also, this is short term in that over time, the control group (that is, nonparticipants) would likely have evolved their actions from one year to the next as conditions change and accounting for these effects is important in determining net savings. As with all approaches discussed in this section, there are pros and cons and the selection of the approach to use and the context in which this choice is made influences these decisions. For example, Tom Eckman of the NW Council indicated that this method may be less controversial in the Northwest because some entities do not have financial incentives tied to estimates of net savings.

NEEP 2012). Although spillover is widely recognized and can be significant, several jurisdictions resist estimating spillover values and including them in the net savings calculations. Market effects values have faced similar challenges.[69]

The NW Council and the RTF have used common practice baselines for energy savings more consistently and longer than any other region or jurisdiction. Much of the RTF work is regional which can help define appropriate markets for both residential and non-residential appliances and equipment. In addition, regional organizations in the northwest (e.g., the Bonneville Power Authority[70] and Northwest Energy Efficiency Alliance) conduct market characterizations for important energy using equipment providing information that the RTF can use to develop these baselines. Finally, the RTF supports the NW Council in the development of a regional power plan every 5 years. The use of the energy efficiency baselines by the RTF are designed to be consistent with the assumptions used in the most recent Power Plan. The RTF has considered the context and needs to be met by its savings estimates and has designed these savings approaches to meet these needs.[71,72] It should also be noted that some entities use other methods discussed in this section to estimate net savings. For example, the Energy Trust of Oregon (ETO), a voting member of RTF, evaluated a smart thermostat energy efficiency program using RED and matching designs (Apex Analytics 2016).

Determining whether a common practice baseline approach provides appropriate savings estimates may depend on a jurisdiction's point of view and how these estimates are used within that jurisdiction. When used as part of a 5-year regional planning process, one point of view might emphasize the estimation of energy savings across five or ten program years. With this perspective, common practice baselines that are re-estimated periodically (as the RTF does) may reflect broad market changes over time. Common practice baselines change over time influenced, in part, by the on-going EE efforts over several years. An alternate view might be applied when looking at incremental resource investments. EE investments that offset other transmission and distribution (e.g., non-wires alternatives) and generation investments may focus on incremental energy savings which may be more appropriate. In this case, the fact that past EE programs may have changed the current EE baseline represents sunk costs and should not be considered economic assessments. If this is the case, only the savings that are incremental and attributable to that year's EE investments should be used. This illustrates how different

[69] To further illustrate, net savings as presented in the findings of EE evaluations are always presented as "net" of something; however, it may be gross savings net free-ridership, or it may be gross savings net free-ridership and spillover, or, in some cases, market effects may be included in the defined net savings estimates. Navigant (2013) found that most jurisdictions defined net savings as "gross savings adjusted only for free-ridership." (The review of net savings methodologies in Navigant [2013a] focused only on C&I programs. Of 38 C&I program evaluations reviewed, 28 estimated net savings as gross savings adjusted for free-ridership only. Three estimated net savings as gross adjusted for free-ridership plus participant spillover, and seven studies adjusted for free-ridership and both participant and nonparticipant spillover. None of the studies attempted to address market effects in addition to the spillover values.)
[70] One good example is the Bonneville Power Administration (2014) market characterization study of non-residential lighting in the northwest.
[71] RTF Guidelines (2015) state "The terms 'net' or 'gross' are intentionally not used to modify the term "savings," as they may conflict with the definition of baseline" in these guidelines.
[72] A presentation by Ms. Jennifer Light, RTF Chair, at the April 2017 Forum meeting presents context around the current practice baselines. https://nwcouncil.app.box.com/v/GuidelinesAprilRTFPres

jurisdictions may have different needs and uses for savings estimates, and how these can drive the approaches used (See NEEP, 2016).

Self-selection bias is a significant concern with common practice baselines. The average action taken in a current market may not be representative of those customers that chose to participate in a specific EE program. A common practice baseline will include a range of equipment with different levels of efficiency. An EE program that allows consumers to select themselves into the program may attract those consumers that comprise that portion of the common practice baseline who would have selected the high-end efficiency equipment. If an EE program attracted those consumers who were predisposed to install the high-efficiency equipment promoted by the program, application of a common practice baseline could overestimate net savings by not accounting for the unique characteristics of those customers. Additionally, to the extent that the program results in nonparticipant spillover, it is not clear how the common practice baseline approach would capture those savings. [73]

4.8.2 Constructing Common Practice Baselines

The theory underlying the definition and pros/cons of common practice baselines can be set out; however, there still is the task of developing these baselines. Developing and maintaining common practice baselines for all the individual measures included in a portfolio of residential and non-residential programs can be a daunting assignment. The RTF in the Northwest built up its library of measure protocols over several years.[74] In addition, the data and information needed for such these multiple baselines can be hard to develop.

SEE Action (2012b) indicates that appropriate common practice baselines can be estimated through surveys of participants and nonparticipants as well as analysis of market data. Discussions with the RTF indicate that they often scan websites of equipment providers to see what types of equipment are currently for sale online. In addition, there are also supporting studies characterizing the markets for energy-using equipment undertaken by other regional entities. A common practice baseline should be based on current sales of equipment and not on the stock of equipment installed. Sales of equipment will represent the current choices of equipment for customers. Sales data can be tough to come by and, even if available, may reflect only parts of the market. Access to sales data will vary by jurisdictions with those jurisdictions that have developed strong connections to equipment manufacturers, suppliers, and trade allies through long-standing EE programs likely to have better access.

When possible, the baseline should be substantiated by actual sales data from retailers and installers, not surveys and anecdotal information.[75] Considerations should include:

1. How much data is required to set a current practice baseline?
    a. Will additional market research and/or studies be required to set current practice baselines?
    b. How will minimum required confidence levels be determined?

[73] This will not be an issue in applications where market-wide sales data are available on standard and energy-efficient equipment, but these data are unavailable in most markets targeted by EE programs.
[74] The list of measures currently addressed by the RTF is available at https://rtf.nwcouncil.org/measures.
[75] Comment from Puget Sound Energy on the proposed baselines for non-residential lighting applications (2016).

2. Subcategories for common practice baselines—by business type, application, region? Regional variation was a significant issue for one IOU in the northwest as pricing and availability were viewed as varying across the service territory.

These considerations seem similar to those discussed in developing initial estimates of savings in Technical Resource Manuals (TRMs) that are developed by many states; however, some argue that common practice baselines with their implicit net savings construction might require a higher level of rigor.

Looking at how the RTF established its "current practice baselines" for two recent measure categories—residential lighting measures and non-residential lighting measures—can provide context for the approaches used for these applications. The non-residential lighting assessment involved the development of a matrix of fifteen applications/replacement types, and six candidate technologies. Not all technologies were appropriate for each application, and 43 efficiencies for incumbent technologies were developed. These were combined with market share estimates of sales to produce a common practice baseline for each of the fifteen applications. Documents available on the RTF website document this process.[76] An additional set of common practice baselines were developed for lighting controls. Also, a dual baseline was used to address installations that represented early replacement and accounted for the remaining useful life of the equipment that was replaced. The value developed for the remaining useful life was one of the more uncertain aspects of the baseline development. This energy savings protocol and baseline assessment represented one of the more complex efforts by the RTF and took more than a year to develop and approve.[77]

Another example of a common-practice type of baseline comes from California's efforts at developing Industry Standard Practices (ISPs).[78] This effort has a narrower focus addressing portable irrigation systems. This PG&E (2016) ISP study of portable irrigation piping systems had the following objectives:

- Evaluate the market trends.
- Determine the common industry practice materials used.
- Understand the barriers for adopting portable PVC (polyvinyl chloride) systems.
- Provide information and guidelines for California utility program developers and stakeholders to consider while developing and managing custom and/or deemed projects.

The data sources and methods used in this ISP included:

[76] The final RTF baseline recommendation for non-residential lighting can be found at: https://rtf.nwcouncil.org/meeting/rtf-meeting-december-6-2016. References to the process used and information from earlier efforts can be found at this site.
[77] Residential lighting energy savings protocols and baseline assessment was another complex effort undertaken. The process, data sources and results of this effort can be found on the RFT website at: https://rtf.nwcouncil.org/meeting/rtf-meeting-march-21-2017 and the supporting baseline presentation at https://nwcouncil.app.box.com/v/20170321ResLightPres.
[78] This example was provided by Dr. Tengfang (Tim) Xu at PG&E Customer Energy Solutions. Revisions to the current approaches for constructing ISP baselines as well as the use of ISP baselines in evaluation including their relationship to net savings are currently being considered (communication with Dr. Xu at PG&E).

1. Development and administration of surveys to customers (growers) and vendors/designers
2. Contacts with customers through emails and phone calls to compliment the surveys
3. Literature research
4. Analysis of the results to determine the market trends and market saturation of the available systems/technologies
5. Highlight critical issues that need to be addressed in custom project development.

Another consideration is that the common practice baseline is essentially a snapshot in time. The common practice baseline will change over time and periodic updates will be needed.[79] The complexity of the update will depend on the program type. For essentially a one-technology program (for example, refrigerator recycling), the update may be straightforward. Updating common practice baselines for a large C&I custom program where many technologies and end uses are impacted may be more difficult. In such cases, it might be more cost effective to focus exclusively on measures that account for the greatest savings. The RTF addresses this by establishing sunset provisions for each energy savings protocol and baseline(s). The sunsets vary across measure categories and are based on how fast the market is perceived to be moving—a market that is evolving rapidly would have a shorter sunset period.

4.8.3 Common Practice Baselines—Summary and Conclusions

Each example of common practice-types of baselines is a best estimate and is subject to uncertainty and potential bias. This is also true with the application of other methods for estimating net savings. Discussions with the engineering teams that developed common practice baselines at the RTF indicated that a practical approach is needed that:

1. Makes use of the best available data for each measure effort
2. Requires the energy savings and baseline engineering team to be open to input from other parties and modifications to initial baseline proposals
3. Develops agreement on the data which helps set the rules for estimating savings and allows for appropriate planning and consistency over time.

Common practice baselines have pros and cons. The decision to use this approach for certain measure categories and programs will depend in part on the jurisdiction's view of the needs and uses of energy savings estimates.

Ridge et al. (2013) make the point that previous EE programs have affected current markets for EE equipment through spillover and market effects. This results in current common practice baselines that are more efficient than they would have been if these past EE programs were not offered. The effect of these past programs is to lower the annual energy use of the measures that constitute the current practice.

---

[79] This is no different than programs evaluated using more traditional methods. The fundamental question is, "What is the shelf life of any evaluation given that many things (e.g., program intervention strategies, technologies promoted, targeted customers, and local and regional economic conditions) can change that would affect the program's ability to deliver net savings?" That is, all evaluations are essentially a snapshot in time.

This argument seems to be partly analytical and partly a policy consideration. Ideally, past evaluations of EE programs should have included all the impacts attributable to the programs, but because spillover and market effects were generally omitted from past evaluations, they have not been counted. The annual energy savings resulting from the use of common practice is lower than it would have been if these past programs were not offered. From this perspective, the use of unadjusted current practice baselines as estimates of net savings seems to be an effort to make up for mistakes in past evaluations (that is, the omission of spillover and market effects that impact the overall market). As a result, a jurisdiction may view savings that accrue today from programs in previous years along with the savings from current programs together comprise a reasonable estimate of EE program impacts over the long term; and, that this best represents the estimate of the overall return on investments in EE.

Another view or position is that each EE program should be evaluated as an incremental investment (that is, a program implemented in 2017 should be evaluated against what is attributable to that investment only—all impacts from prior years' programs are essentially sunk costs and should not be considered). This is an example of where policy and analytic views of net savings estimation are linked.

The bottom line for assessing the common practice baseline approach is the same process that is used in all other methods: (1) understand the construction of the baseline used in the evaluation; and (2) analyze the implications of this baseline against an appropriate counterfactual scenario for that program. The potential uncertainty and magnitude of bias needs to be at least subjectively assessed. Based on this standard approach, decisions can be made about the estimation methods most appropriate for the evaluation of an EE program taking into account jurisdictional priorities and needs.

When an evaluator encounters a jurisdiction that is using a "current practice baseline" method and refers to these savings as net savings, the evaluator should proceed in an internally consistent manner.[80] For example, it is important that the evaluator explain what the utility/agency/regional body is calling gross savings and what, if any, adjustments have been made in the establishment of the baseline to produce a net savings value.

In summary, several jurisdictions are looking toward the use of common practice baselines in their EE evaluation guidelines. As with all methods, there are pros and cons (see Table 7). A

[80] Reviewers of this section have commented that the evaluator might conduct multiple current baseline studies, calculate *ex post* net savings, and calculate a net realization rate to test the robustness of the approach; however, the cost of the analyses becomes a factor. Analyzing the market and different baselines has been presented as useful for understanding EE programs. This view may be most appropriate for jurisdictions that have EE measure and equipment specific data. These data may be limited to certain types of programs, and require a commitment to gathering data at the measure level. Also, before taking this approach, the evaluator might want to make sure that self-selection, nonparticipant spillover, and market effects are not serious sources of bias. If serious bias is suspected, the evaluator could select the baseline from the multiple baseline approaches above as the one that produces the most conservative results; however, there may be little analytic support for this selection. Another suggestion advanced in this newly developed literature is to augment the results using a survey based self-report NTG ratio, but this seems to defeat the purpose of using the common practice baseline method as an *ex ante* method of producing net savings. It increases costs and brings in the issues involved in using appropriate survey methods, and it may thereby reduce some of the advantages claimed for the common practice baseline approach.

potential strength of the common practice baseline approach is its use in upstream and market transformation EE programs. It can be applied market-wide and, unlike randomized trials and quasi-experimental designs, it does not require participants to be identified if appropriate sales data are available. However, this method is susceptible to self-selection bias (that is, the average consumer may not be the type of consumer who participates in the program). It is not clear how this can be addressed, other than by conducting surveys to determine specific characteristics of purchasers of efficient equipment relative to the common practice baseline. However, this survey effort would negate the unique aspects claimed for the common practice baseline approach; i.e., specific consumers who have and have not purchased the high efficiency equipment would need to be identified. This makes this approach more similar to the survey method approaches discussed in Section 4.2.

Table 11. Common Practice Baseline Approach—Summary View of Pros and Cons

Pros	Can help to avoid double counting of free-ridership in circumstances where gross impacts incorporate some net savings factors Can be used in upstream and market transformation programs Can be applied market-wide.
Cons	Self-selection bias is not addressed and methods for addressing self-selection are not readily apparent Does not capture nonparticipant spillover Common practice baselines for measures and technologies will change over time and require updating Determining average market practice has accuracy challenges Approach has been applied in the Pacific Northwest, along with other net savings estimation methods, but is relatively new and still evolving as a general net savings estimation method.

4.9 Top-Down Evaluations (Macroconsumption Models)

Top-down evaluations use macrodata on energy consumption in a model that relates changes in energy consumption to a measure of EE effort (usually expressed as EE expenditures). Top-down evaluation produced macroconsumption metrics (MCMs) in two recent pilot applications in California (see Cadmus 2012a; Demand Research, LLC 2012). The broader literature refers to these as top-down methods, and the MCM notation adopted in the recent California pilot studies refers to the same set of methods and cites top-down studies as background for its pilot work.

To date, this method's application has been somewhat limited. Applications to utility level programs have been limited to pilot studies and the general applicability of these methods has not been demonstrated. Still, the top-down approaches have appeal because they directly address overall net savings. The dependent variable is overall energy use (often expressed as energy use per capita) and this method simply examines the change in energy use resulting from EE efforts. Thus, there is in principle no need to adjust for free-ridership and spillover, or even for market effects, in estimating overall net savings. In addition, the regression analyses provide confidence and precision levels around these estimates. However, there are challenges in estimating the relationship between EE efforts and changes in overall energy consumption, such as the size of the impact isolated by the model.

The development of a model that can measure a 1%–2% change in total energy use annually and is attributable to EE programs requires a reasonably sophisticated structure. For example, the model must have an appropriate lag structure because the impacts from one year's expenditures will occur over several years.[81] In addition, the number of observations and quality of data needed to identify a small effect can be challenging. The data platform needed to support this top-down or MCM model approach requires the following:

- A measure of EE expenditures (or another metric of EE effort for different cross-sections, such as utilities or program administrators)
- Many observations to identify the effects of EE over several years, taking into account the lag structure of EE impacts. As a result, most top-down studies include multi-utility or multi-state efforts that can provide a reasonably large number of cross-sectional areas for the analyses
- Matching demographic and macroeconomic data to utility service areas, or subareas of utilities that are used as observations in the analyses
- High-quality data about energy consumption for each cross-section analyzed.

Questions that evaluators should consider when deciding on the appropriateness or applicability of top-down models are:

- What information will be produced by these top-down models if they are successfully estimated, recognizing that many cross-sections with varying levels of EE investment are needed for estimation?
- How does this information compare to what is produced by other methods?

Top-down models may be useful for:

- Estimating overall average change in energy use from the EE programs for a region. A top-down model that provides a good fit, meets reasonable assumptions, and has acceptable levels of statistical significance can provide information on the average change in overall energy use (or energy use per capita) from overall EE efforts.
- Estimating regional environmental impacts. Aggregate models can be useful in assessing state and regional environmental impacts such as the impact on carbon emissions.
- Providing evidence of estimated energy-savings at a regional level. The model can confirm—at an aggregate level—whether the expected energy savings are actually reflected in the macroconsumption data.
- Estimating overall cost savings from EE programs. Top-down models can also be used to estimate an overall cost savings per kilowatt-hour saved and confirm the efficacy of the overall EE effort.

[81] BC Hydro (2012) demonstrates the importance of the relationship between current expenditures on EE and future savings. It also shows the importance of letting the data determine the most appropriate lag structure as opposed to implementing a fixed structure that acts as a constraint. The estimate of energy savings is influenced by the manner in which lagged effects are handled in the regression model.

Top-down models, however, cannot provide information about:

- Savings produced by specific measures or programs or the impact of an individual program year for the over portfolio
- Where to make additional investments in EE at the program or measure level
- How to improve existing programs
- How to use estimates of free-ridership and spillover to suggest program improvements
- Quality assurance/quality control processes needed for regulatory oversight.

The relative importance jurisdictions and stakeholders place on program-level versus aggregated information will influence decisions to implement these types of evaluation frameworks. Top-down approaches seem complementary to results produced by program-level evaluations; however, there may be concerns about using these methods to replace program-level evaluations. Some view the program-level research as essential in that it helps ensure that the right set of programs comprise the EE portfolio and it is useful in addressing program- and portfolio-specific questions about implementation. Top-down methods and program-level evaluation provide useful, but different, perspectives on the accomplishments of EE efforts.

Cadmus (2012a) reviewed a number of top-down studies that expressed energy consumption as a function of a metric meant to measure EE effort including:

- Parfomak and Lave (1996) used a panel dataset of 39 utilities from 1970 to 1993. The claimed savings by utilities for their C&I programs was used as a proxy for the level of EE effort. The regression analysis was similar to a realization rate regression analysis model, where the coefficient on the claimed utility savings indicated what fraction of those savings could be found in the data. The authors estimated the realization rate for the utility's claimed savings at 99%.
- Auffhammer et al. (2008)—working with data developed by Loughran and Kulick (2004)—used what has become the more traditional formulation. Here, EE effort was expressed in the econometric model as program expenditures reported to the U.S. Energy Information Administration. The authors found that average utility reported savings (2%–3%) fell within the 95% confidence interval for estimated savings. The cost of saved energy was approximately $0.06/kWh.
- Arimura et al. (2011) also used the Energy Information Administration data on program expenditures across 307 U.S. utilities to examine the impacts of EE investments on overall energy consumption.[82] The authors used utility Energy Information Administration data from 1989 to 2006 to determine electricity savings of 1.8% annually and estimated the cost of saved energy at approximately $0.05/kWh.

[82] Arimura et al. (2011) also advance the state of the practice by modeling energy prices and utility EE program expenditures as endogenous and allowing consumption to depend on program expenditures in a flexible way. The literature on top-down models represents sophisticated applications of econometric methods. Problems of endogeneity and autocorrelation with flexible lag structures have become common issues that are addressed by these models.

The California Pilot Project on top-down methods involved two efforts, Cadmus (2012a) and Demand Research, LLC (2012).

In addition to these studies presented below as case studies, the Massachusetts Program Administrators have been piloting top-down studies since 2015. As part of this effort, the evaluation team recently completed a more extensive literature review of top down methodologies applied in studies in both the energy and non-energy sectors. They also conducted in-depth interview with nationally recognized experts in the field of econometrics, macroeconomics, and top down modeling. Some of the methodological considerations examined in the literature review and industry expert interviews included the best approach for handling seasonal weather variations, how top down models should account for the cumulative effects of energy efficiency programs over time, how models should determine the impacts on energy usage for the recession years (2007-2009), the types of fixed effects terms that should be included in a model, as well as considerations from determining error bounds form the model results.

Based on the results of the literature review and expert interviews, the evaluation team provided an assessment of the advantages and disadvantages of different theoretical methodologies and approaches, and provided recommendations for enhancements to future top-down efforts in a February 16, 2017 report to the Massachusetts Program Administrators.[83]

4.9.1 Case Studies on Top-Down Approaches

Example 1: Cadmus California Top-Down Pilot Study

Cadmus used expenditures on EE programs as the level of EE effort in its models. The models were estimated at the utility level for residential and nonresidential energy savings. Cadmus worked with data at the utility level using information from the three investor-owned utilities (IOUs) and from large public utilities in California such as Los Angeles Department of Water and Power and the Sacramento Municipal Utility District. Data were also collected from some small public utilities, but were generally inconsistent.

Several models estimated the relationship between utility energy consumption for residential and nonresidential customer segments and EE expenditures.[84] Overall, it was difficult to obtain significant results across the models. The best model produced significant coefficients on the EE expenditures variable using only data from the three IOUs. To demonstrate the information that can be produced by top-down models, Cadmus developed estimates of savings from EE efforts over a 6-year period and calculated the cost of energy saved. Savings from EE spending from 2005 to 2010 were estimated at 8%, and the cost per kilowatt-hour saved was estimated at $0.05. The results of the Cadmus study indicated savings were within 10% of the net savings reported by California IOUs for the 2006 to 2008 program cycle. The estimates of energy savings and cost per kilowatt-hour saved had large confidence intervals: ±66% on the energy savings

[83] Tetra Tech, NMR Group, and DNV GL (2017). Top-Down Modeling Extended Methods Review. Prepared for the Massachusetts Program Administrators, February 16, 2017.

[84] Cadmus (2012a) did not try to estimate separate models for commercial and industrial consumers because the time series was inconsistent. In some years, commercial sector consumption would increase and industrial consumption would decrease by approximately the same amount. This suggested that there was some switching in the definition on the commercial and industrial rate classes. As a result, the two classes were modeled together.

estimate and more than ±100% on cost per kilowatt-hour saved. The 48 observations in the top-down IOU model resulted in lower precision than studies with much larger sample sizes.

Cadmus did consider disaggregating the data beyond the IOU level to gain more cross-sections for the analysis; however, there was concern about the ability to allocate EE program expenditures to smaller geographic areas. One specific concern was the savings from CFLs. More than 50% of the expected savings were from CFLs and these sales were tracked at point of sale instead of the location where they were used, making it difficult to align the energy consumption and the impact of EE expenditures for smaller geographic areas.

Example 2: Demand Research, LLC, California Top-Down Pilot Study

Demand Research (2012) developed an MCM model working with California utilities and program contractors that disaggregated residential energy use and estimates of residential sector EE efforts into a database of cross-sectional observations at the census tract level. C&I sector energy use and metrics for EE efforts were disaggregated down to the county level. Instead of using energy expenditures, the Demand Research, LLC, study used the utilities' *ex ante* estimates of energy saved by census tract as the metric of residential EE effort. Parfomak and Lave (1996) used a similar approach. For the C&I sectors, county-level data were developed. The independent variable for the EE level of effort in the commercial sector model was a metric related to incentives paid; however, *ex ante* energy savings was used as the metric for EE effort by county for the industrial sector.[85, 86]

The findings from the Demand Research, LLC, study were:

- The residential models estimated by Demand Research, LLC, (2012) showed that higher levels of the EE effort variable resulted in reduced energy use with statistically significant estimates at a 95% confidence interval.
- The commercial sector model produced the expected sign on the EE effort variable, but the results were not statistically significant.
- The industrial sector model did produce statistically significant results for the EE effort variable.
- The residential and C&I sector models produced statewide savings estimates of 7.3% for the 5-year period from 2006 to 2010.

[85] Different metrics for EE level of effort were used in the C&I sector model because the method selected to address endogeneity in the commercial sector model ensured that the EE level of effort variables uncorrelated with the error term.

[86] Considerable work went into creating the census tract databases for the residential model and the county level databases used in the commercial and industrial models. The details can be found in the full study, but as an overview of the effort -- key energy consumption and program tracking data by fuel and segment were inspected prior to modeling for missing values, seemingly erroneous data or outliers, and high and low-end values that might skew the sample statistics or suggest multimodal distributions. Other adjustments to the datasets were made, including the use of a "restricted" commercial sector dataset that included only counties with high *ex ante* energy savings values in this pilot test. Dropping sites from statistical analyses that likely provide no information because the expected savings from those sites are so small is not uncommon. The usual justification is that the total savings number is not likely to be influenced by their exclusion because the expected savings were so small.

- The relative precision for the aggregate savings estimate was ±31% (or a 90% confidence interval of 5.0%–9.5%).
- The estimated statewide savings of 7.3% exceeded the utility *ex ante* estimates of 4.8%.

The aggregate statewide estimate of energy savings across all three sectors was forecasted with reasonable confidence and precision. Looking at the results at one level of disaggregation lower (at the sector level results) shows a high degree of variability. For example:

- The estimated industrial energy savings (all three utilities combined) were about 745% higher than the utilities' *ex ante* values (Demand Research, LLC 2012, p. 36).
- The commercial sector kilowatt-hour savings estimates (all three IOUs combined) were about 27% lower than the utilities' *ex ante* estimates.
- The residential sector savings estimates from the MCM model for Pacific Gas & Electric and San Diego Gas & Electric (Southern California Edison was not estimated) were substantially higher than the utilities' *ex ante* values.

When these sector-level results are aggregated up to a statewide number, the wide discrepancies at the sector level tend to offset each other. It is important to recognize that this was a pilot effort and views will differ about the overall robustness of findings at the sector and statewide levels.

4.9.2 Developing Top-Down Models

Cadmus (2012a) and Demand Research, LLC, (2012) took different paths to developing a top-down MCM model for this California Pilot Study. Both study teams concluded that the work to date indicated this was a potentially useful research path for developing statewide estimates of energy savings attributable to EE policies. In its study report, Cadmus discussed the potential applications of these methods:

- Top-down macroconsumption methods could yield inexpensive[87] estimates of energy savings from utility EE programs and building codes at an aggregate level.
- These methods are attractive because it is possible to produce confidence and precision levels for the net energy savings estimates, which is not as easily accomplished in bottom-up evaluation studies. [88]

[87] Both pilot studies ran into data problems that would have to be overcome in future work and could be costly to address. If the alternative were to build up statewide estimates by doing measure-specific engineering analyses, this aggregate Top-Down approach might be less expensive; however, bottom-up methods performed cost effectively are probably needed for program support, design, and verification of savings at the program level. The issue is whether the incremental information provided by these aggregate studies has a value greater than its cost. That may vary by jurisdiction.

[88] This is a conclusion from the Cadmus (2012a) top-down applications; however, bottom-up approaches also routinely calculate confidence and precision levels for program and portfolio estimates of net savings. The advantage with the top-down approach might be that the confidence and precision levels can be calculated more easily at the aggregate level, because different values for confidence and precision across programs do not have to be combined using assumptions about the covariance across the different distributions from which these values are calculated for each program.

- Top-down studies can be used to verify statewide EE program savings estimates based on bottom-up evaluation by looking at aggregate energy consumption data.
- These methods can be useful in tracking a state's progress in reducing greenhouse gas emissions and developing forecasts of energy savings from future program spending at an aggregate level.

Next steps that might provide additional insights into this top-down application are to: (1) replicate the results of Cadmus and Demand Research, LLC using the datasets already developed; and (2) continue improving the data platform used for these analyses—both studies contained recommendations for improving the data. Violette et al. (2012) discuss the importance of the data platform on which these top-down models are estimated. Other considerations pertain to the sensitivity of the results to model specification (that is, the robustness of the results under a designed set of alternative specifications that are also consistent with the theory and appropriate econometric methods).[89]

Top-down studies cannot entirely replace bottom-up studies (see Table 12 for pros and cons of these methods). As discussed earlier, there is likely a need to have program-level (and some measure-level) assessments to ensure that a program's design will result in a program meeting its specified targets. Moreover, top-down studies are subject to a range of methodological uncertainties not fully captured by the measured precision, just as the bottom-up estimates are. Evaluators should ask, "Does the incremental value of the information produced by the top-down methods exceed the cost of the work?" At the national level, data from an adequate number of cross-sectional observations are more easily available. For state-level studies, more work will be involved in setting up the databases and disaggregating the data into the number of needed cross-sections, which may introduce some error into these observations.[90]

[89] This sensitivity analysis might examine the stability of the estimates under alternative functional forms, inclusion of one or two variables, testing of interaction terms, and tests on subsets of the data.
[90] Violette and Provencher (2012) discuss attenuation bias where the coefficients on independent variable can be biased toward zero due to errors in the measurement of variables. A similar effect is shown in Ridge (1997).

Table 12. Top-Down Evaluations (Macroeconomic Models)—Summary View of Pros and Cons

Pros	Estimates net effects of all programs cumulatively No need to adjust for free-ridership, spillover, or market effects at the aggregate level.
Cons	Methods are not fully developed at the state or regional levels Relies on high-quality energy consumption data and on data regarding EE efforts within each cross-section analyzed Subject to bias and uncertainty due to self-selection, cross-unit spillover, data limitations, and model specification uncertainty Cannot provide savings at the measure, technology, or program level Does not provide information on how to improve program design and implementation processes.

5 Conclusions and Recommendations

A central theme in this chapter is that all decisions have an implicit counterfactual scenario—what would have happened if the decision had not been made. In the context of EE programs, net savings are energy use with the program as compared to a counterfactual that is meant to represent what energy use would have been without the program investments. This chapter does not prescribe specific methods for determining net savings, but rather it presents approaches for assessing attribution and the net impacts of EE programs and discusses the issues affecting the choice of a net savings approach within an evaluation context.

5.1 A Layered Evaluation Approach

It is important that the selected approach be appropriate for the intended audience and present analyses supported by evidence. A well-executed statistical analysis may be a central piece of the evaluation, but it still may not be persuasive to many decision-makers and stakeholders on its own. All approaches should be supported by a narrative discussing why a specific approach was taken, the appropriate interpretation of the findings, and the context for identifying net savings (see historical tracing above). The narrative and analysis should also recognize and indicate the uncertainty in net savings determination. Developing an appropriate narrative often leads to the application of layered methods of analyses.

Studies examining net savings from EE programs may contain both sophisticated quantitative analyses as well as intuitive analyses that show savings that are attributable to the program exist. A compelling part of the narrative can be a simple case study of one or two market participants. A case study can show with a very high degree of internal validity that net savings were obtained, and/or provide examples of NTG factors including free-ridership, spillover, and market effects. An intuitive case study often is a useful first step in a two-part analysis framework to address estimates of net savings. For example:

- **Part 1:** Establish the existence of the effect, possibly using a case study approach. This can include establishing the existence of savings that are attributable to the program. If the focus of the research is on estimating free-ridership or spillover, the first step can involve establishing the existence of these effects. Once existence of an effect is established, the magnitude of the effect needs to be determined. This can be easier when the audience is convinced that the effect exists (i.e., the effect is nonzero), and the logic behind the attribution of the effect is set out.
- **Part 2:** This involves the extrapolation of the findings of the case studies to the more general participant population. Once the logic of the case studies is established, it is often possible to define and apply a statistical model consistent with this logic, or to develop an alternative approach to extrapolate the effect. This approach could include any of the methods discussed in this chapter—survey methods, common practice baselines, market data analyses and comparisons, structured expert surveys, or historical tracing to examine the influence of a program over time.

The framework above for analyzing net savings can be extended to three steps:

1. Perform an initial high internal validity case study to prove the existence of effects.

2. Establish an estimate range. In other words, determine a reasonable lower bound for the impacts and the highest reasonable bound from the evaluation analyses. This provides information about the importance of the studied effect and whether it is a part of net savings or an NTG factor (free-ridership, spillover, or market effect).
3. Perform analyses using the methods presented in this chapter to develop the best estimate of impacts within the established range.[91]

5.2 Selecting the Primary Estimation Method

The selection of appropriate net savings analysis methods will depend in part on the questions that need to be answered by a net savings study. Research issues that have implications for the net savings approach include:

- **RCTs and quasi-experimental designs** employing DiD and regression methods along with RDD and RED designs (discussed in Sections 4.1 and 4.2 of this chapter). These approaches produce estimates of net savings that address free-ridership and participant spillover. Nonparticipant spillover is not directly addressed but can be addressed through surveys of nonparticipants and market effects studies with trade allies.
- **Survey methods** can be used to adjust engineering-based gross savings estimates for free-ridership and participant spillover (discussed in Section 4.3). Nonparticipant spillover can be addressed through surveys of nonparticipants and market effects studies using trade allies.
- **Broader-based methods such as market sales, structured judgment, and historical tracing analyses** can all be used to provide program-specific net savings estimates and address spillover and market effects (discussed in Sections 4.4, 4.5, and 4.7).
- **Deemed or stipulated methods** can be set at the program level (discussed in Section 4.6); however, the applicability from one jurisdiction to another should be considered.
- **Common practice baseline methods** can produce estimates by developing baselines on a program basis (discussed in Section 4.8). This approach may not fully address free-ridership or participant spillover, because it does not account for self-selection bias. Also, it does not directly address nonparticipant spillover. However, as previously noted, nonparticipant spillover can be addressed through surveys of nonparticipants and market effects studies with trade allies. Common practice baseline methods might be viewed as a compromise that balances out over- and underestimated NTG factors in the net savings estimate.
- **Top-down analyses** use aggregate data that represent the overall level of EE effort across all programs, but cannot isolate the effects of a single program or measure (discussed in

[91] In a survey setting, this approach can help the survey respondent consider first the behavior that might result in lower, and then the higher impacts that might have been achieved if the program had not existed. The thought process developed by this three-step construct can help survey respondents produce better estimates of their most likely behavior by first thinking through a construct where the respondent is first asked about factors that would result in a low-range value and then factors that would result in a high-range value.

Section 4.9). Top-down models conceptually address all of the NTG factors—free-ridership, spillover, and market effects.

How can estimates of net savings on a program basis be combined with information about program implementation effectiveness? Approaches that provide estimates of net savings but also include elements that involve gathering information directly from participants, nonparticipants, and trade allies can be useful for improving program performance. For example, some programs are designed to minimize free-ridership to improve overall resource effectiveness and others focus on expanding the magnitude of spillover and market effects. For these programs, specific estimates of free-ridership, spillover, and market effects—particularly if they are provided over a longer time period (every 2 years)—can be used to assess overall program effectiveness.

Can evaluators estimate aggregate net savings from a portfolio of programs? All the estimation approaches presented here, except the top-down analyses, can produce program-specific estimates that evaluators can aggregate up to the portfolio level. Top-down methods are designed to work with aggregate data, particularly at the regional level.

Other factors that influence the selection of appropriate methods will vary by program type, delivery, sector, and maturity. A recent free-ridership and spillover methodology study for the Massachusetts Program Administrators describes the key elements evaluators should consider when choosing a method (Tetra Tech et al. 2011). This study addressed the following factors:

- **Availability of market sales data with a meaningful comparison group**. If market sales data are available on the total sales of both efficient and standard equipment over time, these data are available for the program area, and there is an appropriate comparison area for the appropriate time, total program effects may be estimated based on these data.

  The ideal strategy is to compare the magnitude of the change in sales of energy-efficient equipment relative to the sales of standard equipment in the program area and the comparison area. However, the program tends to produce systematic differences between the program and comparison areas. Therefore, where a program has been operating for a long period of time, it is very difficult to find a comparable comparison area.

- **Homogeneity of the measure and the consumers**. RCTs and quasi-experimental designs work best when there are many similar consumer types and measures. Large custom programs are likely to have fewer projects, so a few (or even one) very large project(s) can have a significant influence on free-ridership or spillover. Therefore, the evaluator should use multiple approaches that allow for a greater focus on the consumers that drive the overall impacts to confirm the findings for that program. Methods based on market data or samples of consumers who are making similar purchase decisions may not apply to programs with custom measures.

- **Likelihood of substantial upstream effects unknown to end-use participants**.[92] If there is a reasonable likelihood of substantial upstream effects that an end-use participant would not know about, then conducting an evaluation by using participating end-user surveys alone will tend to understate the effect of the program (even if consumers answer accurately from their perspectives). These situations require either information for the market as a whole (if the market sales-based approach is viable) or a combination of participant end-user and vendor surveys.
- **Cost/value tradeoffs**. Some methods that provide more credible results are costlier. This cost may be justified for program components that are important to the portfolio, but not for all components. Importance to the portfolio is typically related to the level of spending or savings associated with a program component. However, a component's importance can also depend on future program plans or other "visibility" factors. The systematic assessment of the value of information gained by net savings estimation approaches s compared to the cost of the research is needed to better balance the requests to meet confidence and precision levels for estimates. A target of 90% confidence at ±10% precision simply may not be reasonable for all but the largest programs in a portfolio. This systematic approach can examine the impacts on ratepayers from incorrectly attributing savings to a program. If it is a small program, the impacts on ratepayers will be small as measured with 90% confidence and 15% or 20% precision using a one-tailed test. This can substantively reduce evaluation costs with little impact on the overall equity tradeoffs between ratepayers and utilities.
- **Data quality**. Data quality is a critical factor for all methods. Typical examples of potential limitations to good data quality are: (1) insufficient information in program tracking databases; (2) lack of clear definitions of what is contained in tracking systems (that is, a data dictionary); (3) limitations on the availability of nonparticipant data (including billing data); (4) insufficient number of years of available billing data for participants; and (5) limitations on the availability of market sales data.

5.3 Methods Applicable for Different Conditions

Table 13 lists methods that are suitable for programs with particular features (based on Tetra Tech et al. [2011]). Programs operate in a context and choosing the appropriate evaluation methods requires balancing the advantages and disadvantages of each method. Thus, this table does not list recommendations for a preferred method for a given situation. Rather, it indicates which of the available methods are applicable to programs with specific features. The scales (i.e., low to high) represented in the table for typical cost and complexity are meant to provide an indication of applicability and cost or complexity relative to other methods in Table 13.

[92] For example, the participating customer may not know that the program influence has changed what options are available, lowered the price of the efficient options, and/or increased the sales staff's knowledge and interest in promoting the efficient option.

Table 13. Summary of Methods Applicable to Different Conditions

Net Savings Method	Surveyed Group	Applicability				Typical Cost or Complexity	Special Requirements
		Custom Measures	Measures With Few, Diverse Participants	Large Numbers of Similar Participants	Measures With Substantial Upstream Influence Invisible to Consumers		
RCTs using DiD	None necessary, but could be conducted to help validate the baseline as an appropriate counterfactual scenario	Poor	Poor	Good	Poor	Low	Random assignment of participants and controls
Quasi-experimental design	None necessary but could be conducted to validate or develop better baselines	Poor	Poor	Good	Poor	Low	Matched nonparticipant comparison group
Regression models—Billing data analyses with control variables and Linear Fixed Effects Regression (LFER)	Participating consumers and comparison group consumers	Poor	Poor	Good if there is a valid comparison group	Good if there is a valid comparison group	Low	Need control variables that influence energy use across participants and nonparticipants
Survey based—participants, nonparticipants, and market actors	Participating end users	Good	Good	Good	Poor unless combined with retailer or contractor surveys	Medium	Counterfactual baseline based on survey responses
	Participating and nonparticipating end users	Poor	Poor	Good	Poor unless combined with retailer or contractor surveys	Medium-High	Nonparticipants must be representative of participants
	Retail store managers and contractors	Good	Good	Medium	Good	Medium	
Survey	Retail store managers	Poor	Poor	Good	Good	Low	

Net Savings Method	Surveyed Group	Applicability: Custom Measures	Applicability: Measures With Few, Diverse Participants	Applicability: Large Numbers of Similar Participants	Applicability: Measures With Substantial Upstream Influence Invisible to Consumers	Typical Cost or Complexity	Special Requirements
based - qualitative sales and counterfactual scenario	and contractors						
Structured expert judgment	Experts	Depends on quality of input methods				Low	
Market sales data (cross-sectional studies)	None	Poor	Poor	Good	Good	Low if data are available; high or not possible if data must be developed	Defined market segment
	Manufacturers and regional buyers and distributors	Poor	Poor	Good	Good	Low	
	Retail store managers and contractors	Good	Good	Medium	Good	Medium	
Common practice baseline	Participating and Nonparticipating end-user surveys or market sales data are used	Poor	Poor	Good	Good	Medium to high	Defined market segment
Top-down methods for regional application	None	Requires data on aggregate energy consumption and information on EE effort (expenditures or related program variable) for a large number of cross-sectional observations over a period of time				Depends on the cost of compiling the initial dataset	Aggregate data available on geographic cross-sections

5.4 Planning Net Savings Evaluations—Issues to Be Considered

Evaluation planners should consider several practical issues when planning a net savings evaluation. These include the use of the information, maturity of the program, timing of the study, frequency of net savings estimation, and whether to use multiple approaches. The following bullets summarize these issues:

- **Use of the information**. It is important to consider how the results of the net savings evaluation will be used and the audience for which the evaluation is intended. This can include shareholder incentives, resource plans, program design, and environmental targets (for example, carbon emissions), among other policy goals.[93] The Gross and Net Savings Decision Making Framework and Template (NEEP 2016) provides a standalone template to guide and document key elements that should be considered when making policy decisions in which gross and/or net savings from energy efficiency programs play a role.
- **Maturity of the program**. Almost all programs are assumed to have some free-ridership. The conventional wisdom is that as the program matures (all else equal), observed free-ridership will increase during the study period, but so will spillover and market effects. As a result, it becomes important to test for spillover and market effects as a program matures.
- **Timing of data collection**. To estimate free-ridership, the data should be collected as soon as possible after program participation. This timely measurement minimizes recall bias (Baumgartner 2013), provides apt feedback on program design, and reduces the possibility that the key decision-maker or market actor is no longer available. However, if the objective is to estimate spillover, the ideal time to collect data is at least 1–2 years after program participation, as this allows sufficient time for spillover to occur. Finally, if the objective is to estimate market effects, regular data collection over a period of time is required.
- **Frequency of net savings estimation**. The frequency of net savings or NTG analyses depends on the use of the information. If it is a component of financial incentives for a program administrator, evaluators may need to conduct these studies more frequently. Usually, there is no need to perform detailed net savings studies more than every other year. But, it also depends on the methods used. A statistical analysis of a residential behavioral program can be estimated every year, because persistence is an important issue and study costs are low. NEEP recommends that net savings estimates be made every 2–5 years (Titus and Michals 2008) because several factors can cause estimates of net savings to change over time.
- **Triangulation of NTG approaches**. Using data from multiple sources limits the effects of self-report bias and measurement error (Baumgartner 2013). Using an in-depth methodology with multiple sources also allows evaluators to weight the value of

[93] For example, NEEP (2012) showed that “compared to New England and New York, states in the Mid-Atlantic more commonly use evaluated gross savings for utility regulatory compliance and net savings for program planning and measurement of cost effectiveness. In contrast, New England and New York are more likely to use evaluated net savings; in doing so, they apply NTG values prospectively rather than retrospectively.”

responses from different decision-makers (Megdal et al. 2009). Other data sources often used are: (1) interviews with key decision-makers at the site; (2) project file reviews or project analysis that looks at barriers to project installation, how the project addressed those barriers, and documentation on the participant's decision to go forward with the project; and (3) market data collection, which might include analyses of market sales and shipping data and surveys of market actors (GDS Associates, Inc. et al. 2010; SEE Action 2012b). A recent study conducted for the Massachusetts Program Administrators presents a general approach that can be used by others as they seek to triangulate and integrate the results of two or more net savings studies.[94] The general approach organizes the results from each study in a table that shows findings for each net savings component as well as the qualities or key considerations of each study's results. This approach provides transparency in the factors driving the final net savings estimate.

- **Some evaluation issues are best addressed prior to rolling out a new or revised EE program**. Program design personnel and evaluators should work together in advance of implementing a program design that includes random assignment to discuss the data needed for evaluation that must be collected as part of program implementation.

5.5 Trends and Recommendations in Estimating Net Savings

As discussed in Section 5.4, the choice of approach for estimating net savings will vary depending on the questions asked, the characteristics of the program(s) evaluated, and the ultimate use of the data. However, there are trends in the application of methods:

- The expanded use of informational and behavioral EE programs is leading to a greater use of RCTs and quasi-experimental designs that employ some form of randomization (RDD or RED) to help address self-selection.
- The complexity of programs and the need for assessing market effects is leading to a greater use of informed expert panels and Delphi-types of analyses.
- The need to examine trends in program performance over time and impacts on markets over time is resulting in long-term planning for net savings and NTG factor analyses (for example, regular studies conducted with panel data).
- Net savings studies are increasingly embedded in survey analyses that are also designed to gather information about program implementation effectiveness.
- The value of information from net savings studies is being considered in a more structured manner to help manage evaluation costs (see NEEP 2016). Achieving 90% confidence and 10% precision may be important for a very large EE program, but for a program that is one tenth of the size of the largest program, precision levels are being generated that represent only 1% of the large program. Also, one-tailed tests should be considered, because for some applications, it may be more important to attain a threshold level of net savings with a certain level of confidence than it is to bound the savings estimate both above and below using a two-tailed test. A one-tailed targeted precision level still allows for the calculation of the upper end to the confidence interval Violette

[94] Tetra Tech; NMR Group; DNV GL (2017). *Net-to-Gross Methodology Research* (TXC08)

and Rogers (2012), and there is value to knowing if there was a high likelihood that the target was exceeded by a given amount. The appropriate level of confidence and precision targets are now often reviewed by EE program administrators and regulators to provide fair attribution estimates that minimize risks to ratepayers and to utilities receiving incentives. Navigant (2013a) discusses a loss function approach for assessing the value of information from net savings studies; and information on sampling and the tradeoffs between confidence and precision for EE evaluation can be found in Violette and Rogers (2012) and Khawaja et al. (2013).

It has always been important to consider evaluation options before implementing an EE program or portfolio of programs. However, the importance of planning the types of net savings studies that are needed and the frequency of this measurement prior to program implementation are becoming critically important. Net savings studies embedded in experimental designs that are established prior to consumers becoming program participants allow for:

- The consideration of randomized designs
- The development of the data platform for estimating consumption-based models (including top-down models)
- The collection of information needed for well-run structured expert panel studies.

In conclusion, net savings methodologies continue to evolve and improve over time. No single methodology is appropriate for all programs or measures, and a single methodology is often not the best choice for estimating program or measure net savings. In the end, jurisdictions should design evaluation plans to assess net savings in conjunction with the key stakeholders considering:

- The appropriate schedule for the evaluation effort over time, taking into account the expected value of the information produced versus the cost of the research effort
- Program design and maturity
- The contribution of the program to overall portfolio savings (past, current, planned)
- The evaluation budget, objectives, and value
- Observations and lessons learned from other jurisdictions.

Finally, adequately documenting the methods used and effectively communicating the results of any net savings study are important. The beginning of this chapter presents a framework for persuasive communication.

6 References

AEP (2012). Appendix H, Evaluation of Home Energy Reports, prepared by Navigant. http://www.oracle.com/us/industries/utilities/home-energy-reports-program-3631943.pdf

Agnew, K.; Goldberg, M. (2017). *Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol: Methods for Determining Energy Efficiency Savings for Specific Measure*s. Golden, CO; National Renewable Energy Laboratory. NREL/SR-7A40-68564. http://www.nrel.gov/docs/fy17osti/68564.pdf

Angrist, J.D.; Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

Apex Analytics (2016). Energy Trust of Oregon Smart Thermostat Pilot Evaluation. Prepared for Energy Trust of Oregon. March 1. http://assets.energytrust.org/api/assets/reports/Smart_Thermostat_Pilot_Evaluation-Final_wSR.pdf

Arimura, T.H.; Li, S.; Newell, R.G.; Palmer, K. (2011). *Cost Effectiveness of Electricity Energy Efficiency Programs.* Resources for the Future Discussion Paper 09-48-Rev.

Auffhammer, M.; Blumstein, C.; Fowlie, M. (2008). "Demand-Side Management and Energy Efficiency Revisited." *The Energy Journal* 29(3): 91–103.

Baumgartner, R. (2013). "Survey Design and Implementation Cross-Cutting Protocols for Estimating Gross Savings." Chapter 12 in *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*. www.nrel.gov/docs/fy13osti/53827.pdf.

BC Hydro (2012). *Review of a Top Down Evaluation Study: Rivers & Jaccard*. Prepared for BC Hydro, by Navigant Consulting, Inc., April.

Bodmann, S. (2013). Controlling for Program Participation Self-Selection Bias. Paper and Presentation at the International Energy Program Evaluation Conference.

Bonneville Power Administration (2014). *Northwest Nonresidential Lighting Market Characterization 2010-2012*. Navigant Consulting, Inc. Cadeo Group, https://www.bpa.gov/EE/Utility/research-archive/Documents/Northwest_NonRes_Lighting_Market_Characterization.pdf

Bradlow, E. (1998). "Encouragement Designs: An Approach to Self-Selected Samples in an Experimental Design." *Marketing Letters,* 9(4).

Cadmus (2012a). *CPUC Macro Consumption Metric Pilot Study: Final Report*. Prepared for the California Public Utilities Commission, October.

Cadmus (2012b). *Efficiency Maine Trust Residential Lighting Program Evaluation: Final Report.* Prepared for the Efficiency Maine Trust. www.efficiencymaine.com/docs/Efficiency-Maine-Residential-Lighting-Program-Final-Report_FINAL.pdf

Cadmus; Navigant (2012). *New York Energy $mart^{sm} Products Program Market Characterization and Assessment Evaluation: Final Report.* Prepared for The New York State Energy Research and Development Authority, Victoria Engel-Fowles Project Manager, Project Number 9875, February. See bullet six study in the link: https://library.cee1.org/system/files/library/10360/CEE_Eval_New_York_Energy_Smart_Products_Program_Market_Characterization_and_Assessment_Evaluation_Feb2012.pdf

Cadmus; Navigant Consulting; Opinion Dynamics Corporation (2012). *2012 Residential Heating, Water Heating, and Cooling Equipment Evaluation: Net-to-Gross, Market Effects, and Equipment Replacement Timing*. Prepared for the Electric and Gas Program Administrators of Massachusetts.

Castor, S. (2012). *Fast Feedback Results.* 2011 Final Report prepared for Energy Trust of Oregon. https://www.energytrust.org/wp-content/uploads/2016/12/Fast_Feedback_-_20110.pdf

Commonwealth Edison (2012). *Energy Efficiency / Demand Response Plan: Plan Year 3 (6/1/2010-5/31/2011) Evaluation Report: Home Energy Reports*. Prepared by Navigant Consulting. www.icc.illinois.gov/downloads/public/edocket/323839.pdf

Cook, T.; Scriven, M.; Coryn, C.L.; Evergreen, S.D.H. (2010). "Contemporary Thinking About Causation in Evaluation." *American Journal of Evaluation* 31:105. http://aje.sagepub.com/content/31/1/105

Demand Research, LLC (2012). *Macro Consumption Metrics Pilot Study: Final Report*. Prepared for: California Public Utilities Commission Energy Division, November.

Diamond, A.; Haninmueller, J. (2007). *The Encouragement Design for Program Evaluation*. Harvard University and International Finance Corporation.

DNV GL, ERS (2017). *Massachusetts Commercial/Industrial Baseline Framework.* Prepared for the Massachusetts Program Administrators and Energy Efficiency Advisory Council. http://ma-eeac.org/wordpress/wp-content/uploads/MA-Commercial-and-Industrial-Baseline-Framework-1.pdf

DNV GL (2017). *Impact Evaluation of 2015 Upstream and Residential Downstream Lighting Programs*. Prepared for the California Public Utilities Commission. CALMAC Study ID CPU0152.01.

http://www.calmac.org/publications/2015_LTG4_Impact_Evaluation_Report-FINALES.pdf

Dubin, J.; McFadden, D. (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52(2):345–362

Duflo, E.; Glennerster, R.; Kremer, M. (2007). *Using Randomization in Development Economics Research: A Toolkit.* Centre for Economic Policy Research, Discussion Paper 6059.

Eto, J. (1988). "On Using Degree-days to Account for the Effects of Weather on Annual Energy Use in Office Buildings." *Energy and Building* 12, 113–127.

Eto, J.; Prahl, R.; Schlegal J. (1996). *A Scoping Study on Energy-efficiency Market Transformation by California Utility DSM Programs*. Lawrence Berkeley National Laboratory. http://emp.lbl.gov/sites/all/files/lbnl-39058.pdf .

Fagan, J.; Messenger, M.; Rufo, M.; Lai, P. (2009). "A Meta-Analysis of Net to Gross Estimates in California." Paper presented at the 2009 AESP conference.

Feng, W.; Jun, Y.; Xu, R. (2006). *A Method/Macro Based on Propensity Score and Mahalanobis Distance to Reduce Bias in Treatment Comparison in Observational Study*. Public Health Research, paper pr05. www.lexjansen.com/pharmasug/2006/publichealthresearch/pr05.pdf

Focus on Energy (2017). "Calendar Year 2016 Evaluation Report – Appendix I: Net Savings Analysis Methodologies." Prepared for the Public Service Commission of Wisconsin by Cadmus Group.
https://focusonenergy.com/sites/default/files/Wisconsin%20Focus%20on%20Energy%20CY%202016%20Appendices.pdf

Fowlie, M.; Wolfram, C. (September 2009). *Evaluating the Federal Weatherization Assistance Program Using a Randomized Encouragement Design (RED).* Presented to Environmental Energy Technologies Division (EETD), Lawrence Berkeley National Laboratory https://eta.lbl.gov/news/events/2009/09/11/estimating-the-impacts-of-low-income-weatherization-assistance-using-a-random

Fowlie, M.; Greenstone, M.; Wolfram, C. (June 2015). "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program."
http://e2e.haas.berkeley.edu/pdf/workingpapers/WP020.pdf#page=1

GDS Associates, Inc.; Nexant; Mondre Energy (2010). *Net Savings: An Overview.* RFP 2009. Prepared for the Statewide Evaluator.

Goldberg, M.; Kademan, E. (1995). Is It Net or Not? A Simulation Study of Two Methods. In Energy Program Evaluation: Uses, Methods, and Results, pp. 459–465. Chicago, IL: National Energy Program Evaluation Conference.

Greene, W. (2011). *Econometric Analysis*, 7th Ed., Prentice Hall.

Guo, S.; Fraser, M. (2010). *Propensity Score Analysis: Statistical Methods and Applications*, *SAGE Publications, Inc.* (Note: Chapter 4 provides an updated discussion of the Heckman models to self-selection along with appropriate caveats. This discussion can be found at: www.sagepub.com/upm-data/30234_Chapter4.pdf

Gwet, K.L. (2012). *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters*. Gaithersburg, MD: Advanced Analytics, LLC.

Gwet, K.L. (2010). *Inter-Rater Reliability Using SAS: A Practical Guide for Nominal, Ordinal, and Interval Data*. Gaithersburg, MD: Advanced Analytics, LLC.

Haeri, H.; Khawaja, M.S. (2012). *The Trouble with Freeriders*. Public Utilities Fortnightly. www.cadmusgroup.com/wp-content/uploads/2012/11/Haeri-Khawaja-PUF-TroublewithFreeriders.pdf.

Hall, N; Ladd, D.; Khawaja, M.S. (2013). "Setting Net Energy Impact Baselines: Building Reliable Evaluation Approaches." Paper presented at the 2013 International Energy Program Evaluation Conference, Chicago, IL.

Heckman, J.J. (1979). "Sample Selection Bias as a Specification Error" *Econometrica* 47(1): 153–161.

Ho, D.; Imai, K.; King, G; Stuart, E. (2007). "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Policy Analysis* 15(3); pp.199–236.

Hoefgen, L.; Clendenning, G.; Osman, A.; Keating, K.; Vine, E.; Lee, A.; Stewart, J.; Stoops, J. (2011). "Finding and Counting Market Effects: A New Construction Program Example." Paper presented at the 2011 International Energy Program Evaluation Conference, Chicago, IL.

Hogan, K. (2015). "Getting it Right: Weatherization and Energy Efficiency are Good Investments." Office of Energy Efficiency and Renewable Energy. https://energy.gov/eere/articles/getting-it-right-weatherization-and-energy-efficiency-are-good-investments

Illinois Energy Efficiency Stakeholder Advisory Group (2016). *Illinois Statewide Technical Reference Manual for Energy Efficiency, Version 5.0*. Volume 4: Cross-Cutting Measures and Attachments. February 11, 2016. http://www.ilsag.info/il_trm_version_5.html

Imbens, G.; Lemieux, T. (2010). "Regression Discontinuity Designs: A Guide to Practice." *Journal of Economic Literature* (48); pp. 281–355.

Indianapolis Power and Light (2015). *Demand-Side Management Evaluation Final Report Appendix A: Self-Report Net-to-Gross Evaluation Methodology*. Prepared by the Cadmus Group and Opinion Dynamics. http://www.oracle.com/us/industries/utilities/cadmus-indianapolis-power-light-3697534.pdf

Itron, Inc. (2010). 2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group. Prepared for the California Public Utilities Commission, Energy Division. www.calmac.org/publications/PG%26E_Fab_06-08_Eval_Final_Report.pdf; www.calmac.org/publications/PG%26E_Fab_06-08_Eval_Final_Report_Appendices.pdf

Keating, K. (2009). "Freeridership Borscht: Don't Salt the Soup." Paper presented at the 2009 International Energy Program Evaluation Conference.

KEMA, Inc. (2010). *Final Evaluation Report: Upstream Lighting Program*, Volume 1. CALMAC Study ID: CPU0015.01; Prepared for: California Public Utilities Commission, Energy Division. Prepared by: KEMA, Inc., Prime Contractor: The Cadmus Group, Inc. www.calmac.org/publications/FinalUpstreamLightingEvaluationReport_Vol1_CALMAC_3.pdf

Keneipp, M., Meurice, J.; Alspector, D.; Sutter, M.; Krause, R.; Hines, T. (2011). "Getting MIF'ed: Accounting for Market Effects in Residential New Construction Programs." Paper presented at the 2011 International Energy Program Evaluation Conference. Boston, MA.

Kennedy, P. (2008). *A Guide to Econometrics*, 6th Edition. Wiley-Blackwell, April.

Khawaja, M.S.; Rushton, J.; Keeling, J. (2013). "Sample Design Cross-Cutting Protocols." Chapter 11 in *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measure*s, www.nrel.gov/docs/fy13osti/53827.pdf.

Loughran David S.; Kulick, J. "Demand-Side Management and Energy Efficiency in the United States". The Energy Journal 25 (1), pp. 19-41. 2004.

Mahalanobis, P. (1936). "On the Generalised Distance in Statistics." *Proceedings of the National Institute of Sciences of India 2* (1):49–55. refhub.elsevier.com/S0021-9290(13)00445-4/sbref22

McMenamin, J.S. (2008). *Defining Normal Weather for Energy and Peak Normalization.* Itron, Inc. https://www.itron.com/PublishedContent/Defining%20Normal%20Weather%20for%20Energy%20and%20Peak%20Normalization.pdf.

McKenzie, D. (2009). "Impact Assessments in Finance and Private Sector Development -- What Have We Learned and What Should We Learn?" The World Bank Development Research Group, Policy Research Working Paper 4944, May. See: https://openknowledge.worldbank.org/bitstream/handle/10986/4137/WPS4944.pdf?sequence=1

Megdal, L.; Patil, Y.; Gregoire, C.; Meissner, J.; Parlin, K. (2009). "Feasting at the Ultimate Enhanced Freeridership Salad Bar." Paper presented at the International Energy Program Evaluation Conference, Portland, OR. www.anevaluation.com/pubs/Salad%20Bar%202009%20IEPEC%20paper%205-12-09.pdf.

Messenger, M.; Bharvirkar, R.; Golemboski,B.; Goldman, C.; Schiller, S. (2010). *Review of Evaluation, Measurement and Verification Approaches Used to Estimate the Load Impacts and Effectiveness of Energy Efficiency Programs*. Lawrence Berkeley National Laboratory. http://emp.lbl.gov/sites/all/files/lbnl-3277e.pdf .

Miller, K. (2011). "Cognitive Interviewing." In *Question Evaluation Methods: Contributing to the Science of Data Quality,* pp. 51–76. Jennifer Madans, Kristen Miller, Aaron Maitland, and Gordon Willis (Eds.) Hoboken, NJ: John Wiley & Sons, Inc.

Mort, D. (2013). "Metering Cross-Cutting Protocols." Chapter 9 of *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*." www.nrel.gov/docs/fy13osti/53827.pdf.

Navigant (2013a). *Custom Free Ridership and Participant Spillover Jurisdictional Review*. Prepared for the Sub-Committee of the Ontario Technical Evaluation Committee, May. www.ontarioenergyboard.ca/documents/TEC/Evaluation%20Studies%20and%20Other%20Reports/Ontario%20NTG%20Jurisdictional%20Review%20-%20Final%20Report.pdf.

Navigant (2013b). *Impact and Process Evaluation of the U.S. Department of Energy's Wind Powering America Initiative.* Prepared for: Department of Energy Office of Energy Efficiency and Renewable Energy, Final Report, DOE/EE-0897, May. http://www1.eere.energy.gov/analysis/pdfs/wind_powering_america_evaluation_2013.pdf

Navigant (2016). *ComEd/Citizens Utility Board Energy Saver Program Evaluation Report*. December. http://ilsagfiles.org/SAG_files/Evaluation_Documents/ComEd/ComEd_EPY8_Evaluation_Reports_Final/ComEd_PY8_CUB_Energy_Saver_Evaluation_Report_2016-12-21_final.pdf

NEEP (2012). "*Regional Net Savings Research, Phase 2: Definitions and Treatment of Net and Gross Savings in Energy and Environmental Policy.*" Submitted to the Northeast Energy Efficiency Partnerships: Evaluation, Measurement, and Verification Forum, by NMR Group, Inc. and Research Into Action, December. http://www.neep.org/net-savings-research-phase-2-0
NEEP (2014). *Decision-Framework for Determining Net Savings Approach – Supplemental Document #2 to Principles and Guidance.* Submitted to the Northeast Energy Efficiency Partnerships: Evaluation, Measurement, and Verification Forum, by Johnson Consulting Group, September. http://www.neep.org/sites/default/files/FINAL%20Supplemental%20Document%202%20Decision%20Framework%20for%20Net%20Savings%20Methods%20Johnon%20Issued%20May%202016.pdf

NEEP (2016). *Gross Savings and Net Savings: Principles and Guidance.* Submitted to the Northeast Energy Efficiency Partnerships: Evaluation, Measurement, and Verification Forum, by Dan Violette (Navigant) and Pam Rathbun (Tetra Tech), April. http://www.neep.org/gross-and-net-savings-principles-and-guidance

New York Department of Public Service (2013a). *Guidelines for Calculating the Relative Precision of Program Net Savings Estimates*. Appendix I.

New York Department of Public Service (2013b). *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approach*. Appendix H.

New York Department of Public Service (2012). *Evaluation Plan Guidance for EEPS Program Administrators,* Update #3, Appendix F. Albany, New York.

Nexant (2017). 2016 Load Impact Evaluation of California Statewide Nonresidential Critical Peak Pricing Program and SCTD Commercial Thermostats. Prepared for Joint CA IOUs, *CALMAC ID SDG0307, April.* http://www.calmac.org/AllPubs.asp

NMR Group, Inc. (2017). *MA RLPNC 16-6: Lighting Shelf Stocking.* Report submitted to the Massachusetts Energy Efficiency Program Administrators. http://ma-eeac.org/wordpress/wp-content/uploads/RLPNC-16-6-%E2%80%93-Lighting-Shelf-Stocking-Report.pdf

NMR Group, Inc. and Research Into Action, Inc. (2010). *Net Savings Scoping Paper.* Revised Draft prepared for Northeast Energy Efficiency Partnerships: Evaluation, Measurement, and Verification Forum. http://www.neep.org/net-savings-scoping-paper-1

NMR Group, Inc. and Tetra Tech (2011). *Cross-Cutting Net to Gross Methodology Study for Residential Programs – Suggested Approaches*. Final report prepared for the Massachusetts Program Administrators. http://ma-eeac.org/wordpress/wp-content/uploads/Cross-Cutting-Net-to-Gross-Methodology-Study-for-Residential-Programs-Suggested-Approaches-Final-Report.pdf

NMR Group, Inc.; KEMA; Cadmus Group, Inc.; Tetra Tech (2011). *Massachusetts ENERGY STAR Lighting Program: 2010 Annual Report.* Final Report prepared for Energy Efficiency Advisory Council Consultants, Cape Light Compact, NSTAR, National Grid, Unitil, and Western Massachusetts Electric.
https://www.efis.psc.mo.gov/mpsc/commoncomponents/viewdocument.asp?DocId=935690223

Nonresidential Net-To-Gross Ratio Working Group (2012). *Methodological Framework for Using the Self-Report Approach to Estimating Net-to-Gross Ratios for Nonresidential Customers.* Prepared for the Energy Division, California Public Utilities Commission.

Oak Ridge National Laboratory (1991). *Handbook to DSM Program Evaluation.* Eric Hirst and John Reed, eds., NTIS Pubs., Washington, DC, # ORNL/CON -336, December.

Optimal Energy, Wirtshafter Associates, Inc. (2015). Delaware Evaluation Framework, Prepared for the Department of Natural Resources and Environmental Control.
http://www.dnrec.delaware.gov/energy/information/otherinfo/Documents/EEAC/DRAFT%20Delaware%20Evaluation%20Framework.pdf

Parfomak, P.; Lave, L. (1996). "How Many Kilowatts Are in a Negawatt? Verifying the Ex-Post Estimates of Utility Conservation Impacts at a Regional Level." *Energy Journal* 17 (4).

Peters, J.; McRae, M. (2008). "Free-Ridership Measurement Is Out of Sync with Program Logic…or, We've Got the Structure Built, but What's Its Foundation." In Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings, Washington, DC. www.aceee.org/files/proceedings/2008/data/papers/5_491.pdf.

PG&E (2013). Evaluation of Pacific Gas and Electric Company's Home Energy Report Initiative for the 2010-2012 Program. Prepared by Freeman, Sullivan Co. Available on the CALMAC.org website: www.calmac.org/publications/2012_PGE_OPOWER_Home_Energy_Reports__4-25-2013_CALMAC_ID_PGE0329.01.pdf.

PG&E (October 2016). Industry Standard Practice (ISP) Study of Portable Irrigation Piping Systems. Prepared by BASE Energy, Inc. Prepared for Dr. Tengfang Xu, Industrial, Agricultural and Water Program, Pacific Gas & Electric Company.

PowerStream (July 2016). Advantage Power Pricing Pilot -- Impact and Process Evaluation. Prepared by Navigant. Reference No.: 172240. https://www.powerstream.ca/attachments/PowerStream_APP_Summer_2015_Evaluation.pdf

Prahl, R.; Ridge, R.; Hall, N.; Saxonis, W. (2013). "The Estimation of Spillover: EM&V's Orphan Gets a Home." In Proceedings of the 2013 International Energy Program Evaluation Conference, August.

Provencher, B; Vittetoe-Glinsmann, B.; Dougherty, A.; Randazzo, K.; Moffitt, P., Prahl, R. (2013). *Some Insights on Matching Methods in Estimating Energy Savings for an Opt-In, Behavioral-Based Energy Efficiency Program.* 2013 International Energy Program Evaluation Conference, Chicago.

Provencher, B.; Glinsmann, B. (February 2013). *Evaluation Report: Home Energy Reports – Plan Year 4.* Prepared for Commonwealth Edison Company.

Puget Sound Energy (2012). *Home Energy Reports Program: Three Year Impact, Behavioral and Process Evaluation*. Prepared for: Puget Sound Energy, Prepared by: KEMA, Inc. https://conduitnw.org/_layouts/Conduit/FileHandler.ashx?RID=849

Ridge, R. (1997). *Errors in Variables: A Close Encounter of the Third Kind*. In Proceedings of the 1997 International Energy Program Evaluation Conference. August, Chicago, IL.

Ridge, R.; Baker, M.; Hall, N.; Prahl, R.; Saxonis, W. (2013). "Gross Is Gross and Net Is Net: Simple, Right?" Paper presented at the 2013 International Energy Program Evaluation Conference, Chicago, IL.

Ridge, R.; Willems, P.; Fagan, J.; Randazzo, K. (2009). "The Origins of the Misunderstood and Occasionally Maligned Self-Report Approach to Estimating Net-to-Gross Ratio." Paper presented at the 2009 Energy Program Evaluation Conference, Portland, OR.

Rosenberg, M.; Hoefgen, L. (2009). *Market Effects and Market Transformation: Their Role in Energy Efficiency Program Design and Evaluation.* Prepared for the California Institute for Energy and the Environment and the California Public Utilities Commission Energy Division.

RTF (2012). *Guidelines for the Development and Maintenance of RTF Savings Estimation Methods.* NW Council, Released December 4. http://rtf.nwcouncil.org/subcommittees/deemed/

RTF (December 2015). Roadmap for the Assessment of Energy Efficiency Measures. Meeting Website link: https://rtf.nwcouncil.org/meeting/rtf-meeting-december-8-2015 Roadmap document link: https://nwcouncil.app.box.com/v/OperativeGuidelines-20151208

RTF (April 2017). Guidelines Updates Presentation. By Ms. Jennifer Light. https://nwcouncil.app.box.com/v/GuidelinesAprilRTFPres

Ruegg, R.; Jordan, G. (2007). *Overview of Evaluation Methods for R&D Programs: A Directory of Evaluation Methods Relevant to Technology Development Programs*. Prepared for the U.S Department of Energy: Office of Energy Efficiency and Renewable Energy. https://www1.eere.energy.gov/analysis/pdfs/evaluation_methods_r_and_d.pdf

Rufo, M. (2009). "Evaluation and Performance Incentives: Seeking Paths to (Relatively) Peaceful Coexistence." In Proceedings of the International Energy Program Evaluation Conference, Portland, OR, August.

Sacramento Municipal Utilities District (2011). *Evaluation Report: OPOWER Sacramento Municipal Utilities District (SMUD) Pilot Year2*. Prepared by Navigant. February. http://www.oracle.com/us/industries/utilities/opower-smud-pilot-year-2-3631985.pdf

Sacramento Municipal Utilities District (2013). Load Impact Results from SMUD's Smart Pricing Options Pilot. Prepared by Freeman Sullivan & Co. for Sacramento Municipal Utility District – SMUD contact Ms. Lupe Jimenez

Scriven, M. (2009). "Demythologizing Causation and Evidence." In *What Counts as Credible Evidence in Applied Research and Evaluation Practice*. Stewart I. Donaldson, Christina A. Christie, and Melvin Mark (Eds.). Los Angeles, CA: SAGE Publications.

Sebold, F.D.; Fields, A.; Skumatz, L.; Feldman, S.; Goldberg, M.; Keating, K.; Peters, J. (2001). *A Framework for Planning and Assessing Publicly Funded Energy Efficiency*. www.calmac.org/events/20010301PGE0023ME.pdf

SEE Action (2012a). *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations.* Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkley National Laboratory. https://www4.eere.energy.gov/seeaction/system/files/documents/emv_behaviorbased_eeprograms.pdf

SEE Action (2012b). *Energy Efficiency Program Impact Evaluation Guide*. Prepared by Steven R. Schiller, Schiller Consulting, Inc. See: https://www4.eere.energy.gov/seeaction/system/files/documents/emv_ee_program_impact_guide_0.pdf

Shadish, W.R., Cook, T.D.; Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. New York: Houghton Mifflin Company.

Southern California Edison (2014). *Edison SmartConnect Demand Response and Energy Conservation Annual Report*,. http://www3.sce.com/sscc/law/dis/dbattach5e.nsf/0/98EF18FDB79332B688257CCA007BFA88/$FILE/R1309011%20DR%20OIR%20-%20SCE%20Compliance%20Filing%20Pursuant%20to%20Decision%2008-09-039.pdf

Stryker, A.; Gaffney, K. (2013). "Why the Light Bulb Is No Longer a Textbook Example for Price Elasticity: Results from Choice Experiments and Demand Modeling Research." In Proceedings of the International Energy Program Evaluation Conference, Chicago, IL, August.

Stuart, E.A. (2010). "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25(1):1–21.

TecMarket Works; The Cadmus Group; Opinion Dynamics Corporation; Integral Analytics; Building Metrics; Energy Efficient Homes Midwest (2012). *Indiana Evaluation Framework*. Prepared for the Indiana Demand Side Management Coordination Committee.

Tetra Tech; NMR Group; DNV GL (2017). *Net-to-Gross Methodology Research (TXC08)*. Massachusetts Program Administrators. http://ma-eeac.org/wordpress/wp-content/uploads/Net-to-Gross-Methodology-Research.pdf

Tetra Tech; NMR Group; and DNV GL (2017). *Top-Down Modeling Extended Methods Review*. Massachusetts Program Administrators.

Tetra Tech.; KEMA; NMR Group, Inc. (2011). *Cross-Cutting (C&I) Free Ridership and Spillover Methodology Study Final Report.* Massachusetts Program Administrators. http://ma-eeac.org/wordpress/wp-content/uploads/Massachusetts-PAs-Cross-Cutting-CI-Free-ridership-and-Spillover-Methodology-Study.pdf

Titus, E.; Michals, J. (2008). "Debating Net Versus Gross Impacts in the Northeast: Policy and Program Perspectives." ACEEE Summer Study on Energy Efficiency in Buildings, (5): 312–323. https://www.aceee.org/files/proceedings/2008/data/papers/5_429.pdf

U.S. Department of Energy (2010). *Guidance Document #7: Topic: Design and Implementation of Program Evaluations that utilize Randomized Experimental Approaches.* Smart Grid Investment Grant Technical Advisory Group, November. https://www.smartgrid.gov/files/cbs_guidance_doc_7_randomized_experimental_approaches.pdf

Violette, D. (2013). "Persistence and Other Evaluation Issues Cross-Cutting Protocols." Chapter 13 in *Uniform Methods Project for Determining Energy Efficiency Program Savings for Specific Measures*. NREL/SR-7A30-53827, April. See: www1.eere.energy.gov/wip/pdfs/53827-13.pdf

Violette, D.; Barkett, B.; Schare, S.; Skumatz, L.; Dimetrosky, S. (2005). *Commercial/Industrial Performance Program (CIPP) Market Characterization, Market Assessment And Causality Evaluation.* Prepared for NYSERDA, Jennifer Ellefsen, Project Number 7721, March.

Violette, D.; Brakken, R.; Schon, A.; Greer, J. (1993). *Statistically-Adjusted Engineering Estimates: What Can the Evaluation Analyst Do About the Engineering Side Of the Analysis?* Published in the *Proceedings of the 1993 International Energy Program Evaluation Conference (IEPEC)*.

Violette, D.; Keneipp, M.; Ozog, M., Stern, F. (1991). *Impact Evaluation of Demand-Side Management Programs — Volume 1: A Guide to Current Practice*. Electric Power Research Institute Pubs., Palo Alto, CA, #EPRI CU-7179, February.

Violette, D.; Ozog M.; Cooney, K. (2003), Retrospective Assessment of the Northwest Energy Efficiency Alliance -- Findings and Report. *Prepared for:* Northwest Energy Efficiency Alliance, Ad Hoc Retrospective Committee, December 8. See: http://www.theboc.info/pdf/Eval-BOC_SummittBlue_NEEA_2003.pdf

Violette, D.; Provencher, B. (2012). *Review of a Top Down Evaluation Study: Rivers & Jaccard (2011).* Prepared for BC Hydro, Navigant Consulting, Inc., April.

Violette, D.; Rogers, B. (2012). *A Sampling Methodology for Custom C&I Programs*. Prepared by Navigant Consulting, Inc., prepared for the Ontario Natural Gas Technical Evaluation Committee, Ontario Energy Board, November. Report available at Ontario Energy Board website: https://www.researchgate.net/publication/279531668_MAIN_REPORT_-_A_Sampling_Methodology_for_Custom_CI_Programs_Technical_Evaluation_Committee

Violette, D.M.; Provencher, B.; Sulyma, I. (2012). "Assessing Bottom-Up and Top-Down Approaches for Assessing DSM Programs and Efforts." In International Energy Program Evaluation Conference Proceedings, Rome, June.

Weiss, C. (1997). "Theory-Based Evaluation: Past, Present, and Future" Special Issue: Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods, New Directions in Evaluation, Volume 1997, Issue 76.

West, S. (2008). "Alternatives to the Randomized Controlled Trial." *American Journal of Public Health* 98(8). http://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2007.124446

Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA, MIT Press Ltd.

Xcel Energy (2016). 2017-19 RE-TOU & RD-TDR Study and Evaluation Plan. November 15 Filing. http://www.dora.state.co.us/pls/efi/EFI.Show_Filing?p_session_id=&p_fil=G_726523

7 Appendix A: Price Elasticity Studies as a Component of Upstream Lighting Net Savings

Studies of upstream changes in the price for residential lighting products have received attention as a way to complement surveys with market actors, or even replace these surveys with econometric models. The way in which price can be viewed as a driver of program savings and the importance of other program components is discussed in Stryker and Gaffney (2013).

Price elasticity studies are currently being applied in several jurisdictions. To date, these studies have focused on residential lighting products and, within that category, mostly on CFL sales. For example, Cadmus (2012b, 2013) and KEMA (2010) tested several different methods for estimating the increase in CFL sales resulting from a program-induced price reduction caused by program activities (markdowns negotiated with retailers and coupons). These two approaches are outlined below.[95]

Cadmus (2012b) examined Efficiency Maine's residential lighting program and Cadmus (2013) examined Wisconsin's Focus on Energy residential lighting program. Both studies used a price elasticity approach. These two studies estimated expected bulb purchases (and associated savings) at prices offered under the program and then the purchases that would have occurred at original retail prices. The difference between these two values was viewed as net savings in this study.

Cadmus (2012b, 2013) used a single equation regression model where the quantity of CFLs purchased was a function of the price of CFLs and a select set of other independent variables. The data used to estimate this equation included package and bulb sales for each retailer, by model number and by week. The dataset does not include information about the consumers who purchased the CFLs, but does contain information about quantities of CFLs sold and retailer prices. Consumer variables desirable in a demand equation would include income and education, but often these variables are not available in the retailers' sales tracking systems.

A regression was estimated relating quantities of CFLs sold by retailer to the price of CFLs that week for each retailer. Other factors such as promotional events were considered in determining consumer purchases. Programmatic factors such as labeling and information dissemination are pervasive throughout the lighting programs and, while potentially important, could not be addressed due to lack of variation across consumer purchases.

These two studies showed an increase in the sales of CFLs as prices decreased due to markdowns negotiated with retailers and discount coupons provided to consumers. The second step of the approach involved estimating what the sales would have been at the higher prices that would have prevailed without the program (that is, the counterfactual scenario).

[95] Both Cadmus and KEMA (now DNV GL) have completed more recent studies using price elasticity approaches for upstream lighting programs. Each incorporates several new features, but the constructs are similar to those discussed in this section. Updated citations for more recent applications are Focus on Energy (2017) and DNV GL (2017).

Considerable effort was made in these price elasticity studies to control for factors other than price that might also affect CFL sales, but it is difficult to show that any method is free of bias. In the case of the Efficiency Maine lighting program, there were three components to the program. Two were linked to price (markdowns and coupons) and a third was linked to overall participation in the Appliance Rebate Program, "with Appliance Rebate Program participants electing to receive a free six-pack of CFL bulbs, via a check-off on the Appliance Rebate Program application form." The third part of the program would have provided CFLs at essentially no cost and it is not clear how this would have factored into the analysis.

Cadmus (2012b, 2013) present several general caveats to the demand equation approach used in the study. First, it acknowledged that "this estimation method has rarely been used in upstream lighting program evaluations as such data generally have been unavailable. As Efficiency Maine … tracked these data and shared them for this evaluation, Cadmus found such econometric demand estimation provided the best method for estimating the program's free-ridership." Second, Cadmus (2013) indicates that it "will continue to look for alternative methods to calculate net-to-gross," and that "the model used for the … 2012 evaluation does not account for spillover."[96]

KEMA (2010) used price variables to estimate net savings in an upstream lighting study. This study had the benefit of a sizeable data collection effort that included consumer surveys. As part of the in-store consumer intercept research, brief interviews were conducted with shoppers who had just made a lighting purchase (revealed preference) as well as "stated preference" surveys with other consumers recruited randomly. Intercept surveys were conducted with 1,463 customers across 378 stores.

KEMA (2010) used three primary types of methods for estimating net savings:

- Supplier and consumer self-report methods
- Econometric models
- Total sales (market-based) approach.

Among the econometric modeling efforts, four econometric models were used:

- Pricing (price formation model)
- Conjoint elasticity
- Revealed preference purchase
- Stated preference purchaser elasticity.

The first two econometric methods—price formation and the conjoint elasticity model—were both needed to produce a net savings estimate. Revealed preference and stated preference models

[96] Cadmus (2012) indicates that spillover is not addressed in this study; however, looking at the overall change in sales in a market caused by price elasticity, has included spillover elements in other studies that use a similar price elasticity approach.

can produce net savings directly. As a result, there were four econometric models, but only three approaches for estimating net savings.

The price formation model estimates the percentage reduction in CFL prices that resulted from program incentives. This is combined with the conjoint analysis, which estimated the corresponding percentage increase in market share/sales that result from a price decrease. This allowed the net savings to be calculated by combining the findings from the pricing study with the conjoint demand elasticity study—in other words, the program induced reduction in prices from the pricing study multiplied by the estimate of change in sales caused by a lower price from the conjoint study.

KEMA (2010) revealed a preference for store intercepts to survey customers that made actual CFL purchases. These customers were asked to indicate how many CFLs they would have bought compared to their actual purchases at double the price they actually paid. Response categories were: (1) the same amount, (2) fewer, and (3) none. Although still based on hypothetical, self-reported responses, the revealed preference respondents may be a more reliable sample because they just made an active purchase decision. However, revealed preference respondents may be somewhat unlikely to indicate they would have paid more for what they just purchased. KEMA (2010) used a random survey of customers, including customers who did not actually purchase a CFL. KEMA (2010) states that the magnitude of the potential bias across these two methods is unknown, "but it is likely that NTG ratio estimates from stated preference respondents are biased downward and NTG ratio estimates from revealed preference respondents are biased upward."

The revealed preference model allowed KEMA to use the store-intercept survey data to model CFL purchase rates with and without program effects. This model was based on a logistic regression to model the probability of buying a CFL rather than an "equivalent" non-CFL as a function of price, displays, customer characteristics, and bulb characteristics, by channel. The fitted models were evaluated under program and non-program conditions. For each channel, the difference between the probability of purchasing CFLs under the program condition and that under the non-program condition was the program-attributable CFL sales share.

In summary, the price elasticity studies completed to date have been limited to residential lighting programs. Cadmus (2012b, 2013) developed a demand model specification based on an examination of alternative specifications. KEMA (2010) developed several approaches for examining the change in CFLs sold as a function of program-induced lower prices. KEMA (2010) concluded that from the econometric approaches, the revealed preference model was the preferred approach. It should be noted that these approaches focus on free-ridership and do not address spillover or longer-term market effects. Currently, several evaluations are using the price-elasticity method to estimate net savings from residential lighting. An expanded literature will likely provide additional confidence in this method for addressing free-ridership from upstream lighting programs, and possibly an expansion of this method to other residential product programs.

Custom Free Ridership and Participant Spillover Jurisdictional Review

Prepared for:
Sub-Committee of the Ontario
Technical Evaluation Committee

May 29, 2013

Prepared by:
Debbie Brannan, Dan Violette, Ken Seiden,
Jane Hummer, and Jeff Erickson

Navigant Consulting, Inc.
1375 Walnut Street, Suite 200
Boulder, CO 80302
303.728.2500
www.navigant.com

NAVIGANT

Acknowledgements

The authors wish to acknowledge Melinda Clarke and Leslie Kulperger of Union Gas Limited and Ravi Sigurdson and Marc Hull-Jacquin of Enbridge Gas Distribution for their guidance, assistance, and support of this work. The authors also wish to acknowledge Jay Shepherd and Bob Wirtshafter, independent members of the sub-committee of Ontario's Technical Evaluation Committee who provided thoughtful feedback and direction on earlier drafts of this report. The authors appreciate the opportunity to work with such a knowledgeable team.

NAVIGANT

Table of Contents

NAVIGANT

Tables

Figures

Executive Summary

Union Gas Limited (Union) and Enbridge Gas Distribution (Enbridge) have delivered Demand Side Management (DSM) initiatives since 1997 and 1995, respectively, including programs that involve custom projects in the commercial and industrial (C&I) sectors. In 2007-2008, Summit Blue Consulting (now part of Navigant's Energy Practice) conducted the first attribution study of Union and Enbridge's custom C&I programs to evaluate free ridership (FR) and spillover effects. After the study, the Ontario Energy Board (OEB) approved the FR adjustment, but did not approve the spillover factor. Since that time, there have been a host of program environment changes, including economic conditions, energy prices, advances in technology, as well as changes in the design and delivery of the custom programs. As a result, Ontario's Technical Evaluation Committee (TEC) is prioritizing updates to FR and spillover adjustment factors as part of its mandate.

This report provides information to support a sub-committee of Ontario's TEC in its deliberations on the appropriate approach to Net-to-Gross (NTG) values in Ontario. Through a jurisdictional review of the approach to net savings, and a review of researched NTG values for programs comparable to Union and Enbridge's custom C&I gas programs, Navigant provides an assessment of the various approaches to NTG.

ES 1. Report Objectives

There are a range of options for NTG that could be adopted for natural gas DSM programs in Ontario, from transferring NTG values from similar jurisdictions and programs to conducting research to estimate a NTG value.

The objective of this report is to provide information to assist the TEC sub-committee in their determination on the appropriate approach to NTG for DSM programs in Ontario, and not to provide a specific recommendation. While this report is not comprehensive in addressing all potential considerations, such as other benefits of accurate (costs of inaccurate) NTG values, it provides important information relevant to the discussion. In addition to summarizing the regulatory and methodological approach taken by other jurisdictions, and summarizing NTG values for programs with characteristics similar to Union and Enbridge's custom C&I programs, Navigant provides insight into the risks associated with inaccurate NTG values and the approximate cost of mitigating those risks.

ES 2. Key Findings

To achieve the objective of this report, Navigant (1) reviewed the approach to net savings across a wide array of jurisdictions in the United States and Canada to identify trends in the regulatory and methodological approach to net savings, (2) conducted a review of researched NTG values of non-residential gas programs in selected jurisdictions, and (3) conducted a decision analysis to assess the options for NTG. Key findings are presented for each of these.

Approach to Net Savings

Navigant conducted research to provide a summary of the regulatory and methodological approach to net savings adopted by jurisdictions across North America. In total, Navigant reviewed the approach to net savings taken by 42 jurisdictions across North America, representing the vast majority of jurisdictions with ratepayer-funded energy efficiency programs.

The majority of jurisdictions with ratepayer funded energy efficiency programs conduct NTG research, though only half adjust gross savings based on research. While there appears to be a trend towards considering participant and non-participant spillover in NTG research in recent years, the majority of research only includes FR adjustments. Both FR and spillover are most commonly estimated through a self-report (participant survey) approach, though econometric methods (e.g., billing analysis) and market share modeling approaches are occasionally used.

Navigant also researched whether jurisdictions offer utility performance incentives for meeting their savings goals. U.S. states that provide a performance incentive mechanism for utilities or program administrators are more likely to make deemed or researched NTG adjustments.

Researched NTG Values in Selected Jurisdictions

Navigant reviewed a total of 19 documents that conducted NTG research of non-residential gas programs covering nine jurisdictions in North America, including: California, Colorado, Massachusetts, Minnesota, New Jersey, New Mexico, Oregon, Washington, and Wisconsin. Within these 19 documents, 38 distinct NTG values were reported.

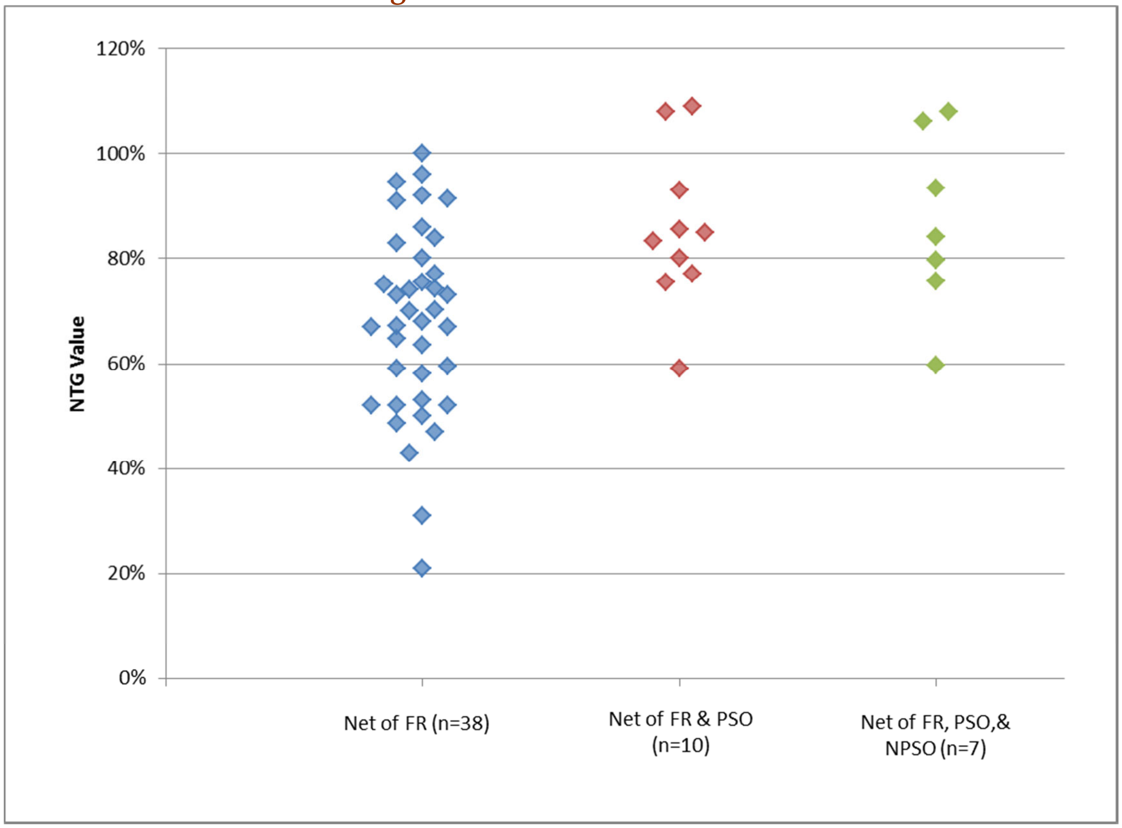
Different formulations of NTG values are presented, with each including or excluding different NTG factors. In particular, the following NTG values are presented:

- Net-of-free ridership = 1- FR,
- Net-of-free ridership and participant spillover = 1 – FR + PSO, and
- Net-of-free ridership and all spillover = 1- FR + PSO + NPSO
  (Note: NPSO is non-participant spillover)

This approach conveys information on NTG values based on the common definitions across the studies, and avoids inappropriate comparisons that could result from comparing the studies' reported NTG values when they include different components.

A review of researched net-of-free ridership values for non-residential gas programs exhibits a wide dispersion (21% to 100%) with a slight "clustering" of values between 40% and 90%, as shown in Figure ES-1. The average net-of-free ridership value is 68%. As expected, NTG values are larger when considering spillover. Average net-of-free ridership & PSO value is 86% and average net-of-free ridership & spillover value is 87%, suggesting that NPSO is small for non-residential gas programs.

Figure ES-1. NTG Values



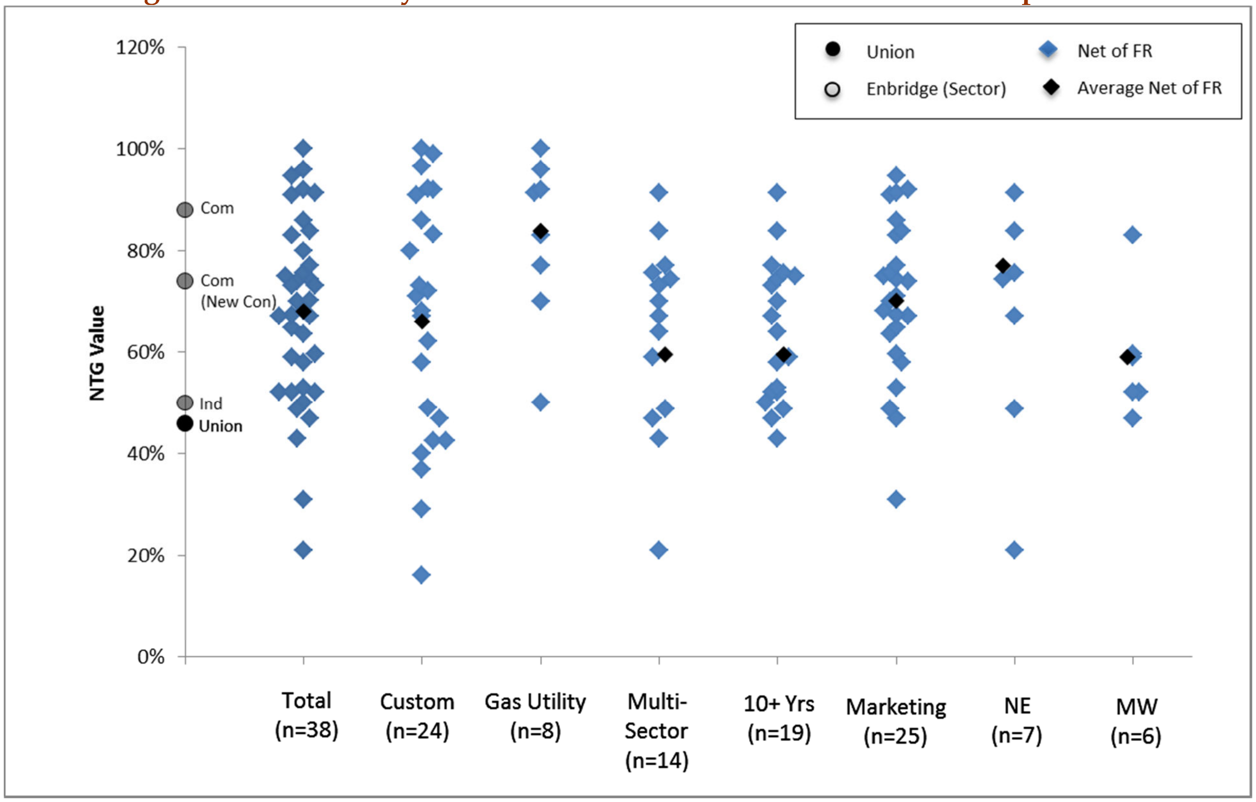
Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG values (program-utility-year combinations) reported in the 19 studies.

To provide additional context Navigant reviewed NTG values by study, program year and region and found that the variation in NTG values did not appear to be driven by the program evaluator, program year, or region. Navigant also examined whether variation in NTG values resulted from differences in the analytic rigor of the methodology (all used self-reports), using enhanced self-report methods in the form of trade ally feedback as a proxy. Free ridership values appeared lower with the inclusion of trade ally feedback. Finally, Navigant compared electric NTG values to gas NTG values for studies that reported both values and found that gas NTG values exhibited a wider dispersion.

Navigant also reviewed researched NTG values based on specific program characteristics: program type, customer segment, utility-type, program maturity, and program marketing strategy. Trends in NTG values are less defined and should be interpreted with caution due to the small sample sizes. Nevertheless, some trends emerged: NTG values for custom programs exhibited a wider dispersion than programs offer prescriptive incentives or both, programs offered by gas-only utilities appear to have lower FR than programs offered by combination utilities, and FR appears to be greater with program maturity.

Figure ES-2 presents the net-of-free ridership values for program characteristics that are most similar to Union and Enbridge's custom C&I programs. In addition, Union and Enbridge's current NTG values, based on the 2007-2008 research conducted by Navigant (formerly Summit Blue Consulting) are presented. Note that Union currently uses one NTG value for C&I custom programs while Enbridge uses sector-specific NTG values.

Figure ES-2. Summary of Relevant Researched Net-of-Free Ridership Values



Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG values (program-utility-year combinations).

Both Union and Enbridge's current NTG values are within the range of researched values. Union's NTG value is below the average value. Enbridge's NTG value for the commercial sector is above the average value while the NTG value for the industrial sector is below the average value.

Assessing Options for NTG

Gross savings can usually be estimated quite accurately, however, estimating net savings poses greater challenges. Given the uncertainty around any NTG value, Navigant applied a Decision Analysis approach for organizing information around alternative approaches to setting NTG values.

There are a number of benefits resulting from more precise NTG values, including the ability to improve program design and implementation, more accurate utility incentive payments, and the ability to consider energy savings as a resource. Navigant conducted a value of information

(VIF) analysis on the second benefit, incentive payments, as the benefit/cost of improved information can be easily quantified.

To support the VIF analysis, Union and Enbridge conducted a sensitivity analysis of utility incentive payments resulting from their custom programs, using a +/- 10 percentage point margin of error on the custom programs NTG values. This analysis revealed that improving the precision of custom NTG values has a sizable impact on incentive payments. Table ES-1 and Table ES-2 present a value of information analysis for Union and Enbridge respectively at targeted net savings.

Table ES-1. Value of Information Assessment for Union

	NTG Value for Custom Programs		Incentives	Change in Incentives
Base Case:	**Current NTG** NTG = 0.46	→	Incentives = $2.73 M	
Scenario 1:	**Higher True NTG** NTG = 0.56	→	Incentives = $5.63 M	**(+$2.90 M)**
Scenario 2:	**Lower True NTG** NTG = 0.36	→	Incentives = $0.8 M	**(-$1.93 M)**

Source: Sensitivity analysis provided by Union.

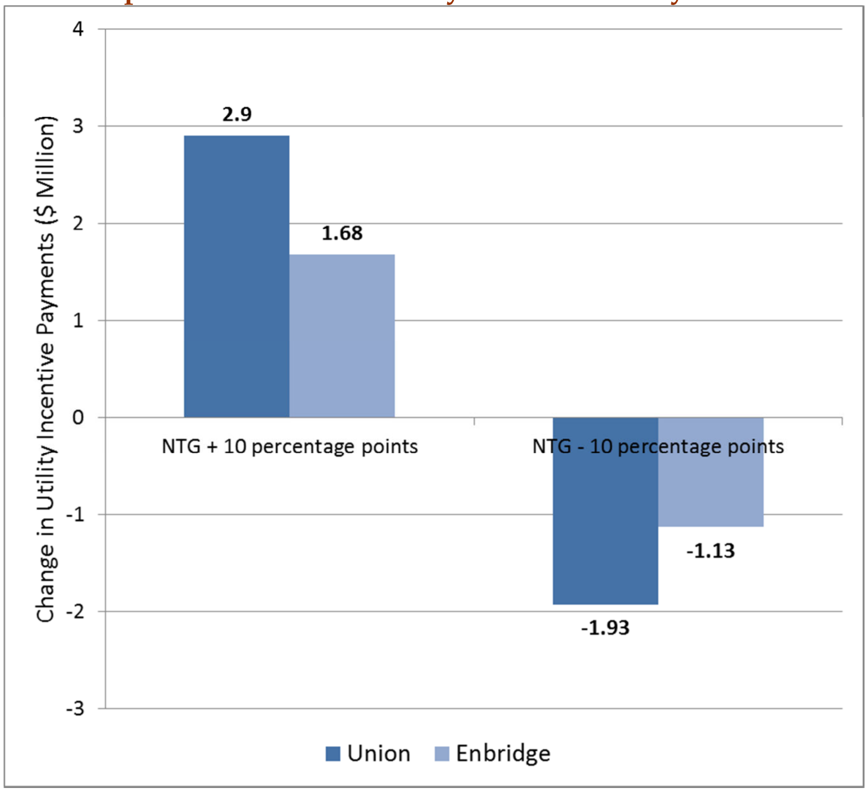
Table ES-2. Value of Information Assessment for Enbridge

	NTG Value for Custom Programs		Incentives	Change in Incentives
Base Case:	**Current NTG by Program** Commercial = 0.80 Commercial New Construction = 0.74 Industrial = 0.50	→	Incentives = $2.58 M	
Scenario 1:	**Higher True NTG** Commercial = 0.90 Commercial New Construction = 0.84 Industrial = 0.60	→	Incentives = $4.26 M	**(+$1.68 M)**
Scenario 2:	**Lower True NTG** Commercial = 0.70 Commercial New Construction = 0.64 Industrial = 0.40	→	Incentives = $1.45 M	**(-$1.13 M)**

Source: Sensitivity analysis provided by Enbridge.

The penalty for assuming a NTG value that is +/- 10 percentage points different from the actual NTG value is roughly $1 to $3 million in utility incentive payments, as shown in Figure ES-3. If the cost of revising the NTG values is less than $0.5 million then revising the values *could be judged to be warranted* assuming NTG research could reduce the margin of error by one-half (i.e., the range of the likely true NTG values).

Figure ES-3. Comparison of the Sensitivity of Incentive Payments to NTG Values



Source: Sensitivity analyses provided by Union and Enbridge.

Navigant provides a brief review of five general approaches to NTG, providing an estimate of the improved precision of the NTG value and the approximate cost per utility (Table ES-3). Alternate NTG approaches could improve the precision of NTG values by approximately 50% at an approximate cost of $0.25 - $0.50 million per utility.

Table ES-3. Ability of NTG Approaches to Produce More Precise NTG Values

General NTG Approach	Estimated Improved Precision (or Reduced Range) of NTG Value	Cost of NTG Approach per Utility (approximate)
Transfer NTG Values from Other Research	Little change	$3 – 5k
Adjust NTG Values based on Program Factors	Little change	$5 – 10k
Align NTG Values using Limited Primary Data	3 percentage points	$100 – 200k
Full NTG Research Study – After Program Year	5 percentage points	$250 – 500k
Integrated/Fast Feedback NTG Estimation	5 percentage points	$250 – 500k

Source: Navigant analysis.

1. Introduction

This report provides information to support the sub-committee of Ontario's TEC in its deliberations on the appropriate approach to NTG values in Ontario. Through a jurisdictional review of the approach to net savings, and a review of researched NTG values for programs comparable to Union and Enbridge custom C&I gas programs, Navigant provides an assessment of the various approaches to NTG.

1.1 Background

Union and Enbridge have delivered Demand Side Management (DSM) initiatives since 1997 and 1995, respectively, including programs that involve custom projects in the C&I sectors. Custom projects cover opportunities where savings are linked to unique end uses and technologies. The DSM portfolio for both utilities includes several hundred custom projects annually. Union and Enbridge DSM activities are regulated by the OEB.

In June, 2011, Union and Enbridge entered into a new DSM regulatory framework. In addition to filing comprehensive, multiyear program plans, Union and Enbridge established Terms of Reference (ToR) for engaging stakeholders. The ToR established engagement processes, and included the creation of a common TEC for both gas utilities. The goal of the TEC is to "establish DSM technical and evaluation standards for measuring the impact of natural gas DSM programs in Ontario."[1]

In 2007-2008, Navigant (formerly Summit Blue Consulting) conducted the first attribution study of Union and Enbridge's custom C&I programs to evaluate FR and spillover effects.[2] The OEB approved the FR adjustment, but did not approve the spillover factor. Since that time, there have been a host of program environment changes, including economic conditions, energy prices, advances in technology, as well as changes in the design and delivery of the custom programs. As a result, the TEC is prioritizing updates to FR and spillover adjustment factors as part of its mandate.

1.2 Report Objective

There are a range of options for addressing net savings that could be adopted for natural gas DSM programs in Ontario, from deeming a NTG value to conducting research to estimate a NTG value. The objective of this report is to provide information to assist the TEC sub-committee in their deliberations on appropriate approaches for developing an NTG value for these programs. This report is not meant to provide a specific recommendation, but rather to

---

[1] *2012 Custom Free Ridership and Participant Spillover Jurisdictional Review Request for Proposal*, Ontario Natural Gas Technical Evaluation Committee, October 29, 2012.

[2] *Source:* Summit Blue Consulting. 2008. *Custom Projects Attribution Study.* Union Gas Limited and Enbridge Gas Distribution, October 27, 2008.

provide information on the range of approaches to assist the TEC sub-committee in making their determination.

The steps taken to achieve this objective include the following:

- Understand the portfolio of Union and Enbridge's custom C&I gas programs (Section 3)
- Review the approach to net savings across a wide array of jurisdictions in the United States and Canada to identify trends in the regulatory and methodological approach to net savings (Section 4)
- Conduct a review of researched NTG values of non-residential gas programs in selected jurisdictions (Section 5)
- Conduct a decision analysis to assess the options for NTG (Section 0)

2. Methodology

This section describes the methodology Navigant employed to provide information to assist the TEC sub-committee in their deliberations on the appropriate approach to NTG for custom natural gas DSM programs in Ontario. The sub-sections that follow discuss the four distinct tasks conducted by Navigant:

- Reviews of the custom C&I natural gas programs,
- Summary of research methods and regulatory approaches to net savings,
- Review of researched NTG values in selected jurisdictions, and
- Assessing options for updating NTG values for these programs.

2.1 Union and Enbridge Programs

To develop an understanding of the portfolio of Union and Enbridge's custom C&I gas programs, Navigant conducted a review of the following:

- Description of programs included in the *2012 Custom Free Ridership and Participant Spillover Jurisdictional Review* request for proposal, and
- Union and Enbridge program websites.

Union and Enbridge also provided additional information on features of program design and implementation as requested by Navigant.

2.2 Approach to Net Savings

Navigant conducted research to provide a summary of the regulatory and methodological approach to net savings adopted by jurisdictions across North America, as well as whether jurisdictions offer utility performance incentives for meeting their savings goals. The research methodology included a review of:

- Utility websites,
- Regulatory agency websites,
- Websites of research/advocacy groups such as the Regulatory Assistance Project (RAP), American Council for an Energy-Efficiency Economy (ACEEE), Consortium for Energy Efficiency (CEE), and the Edison Foundation, and
- Studies that previously surveyed the approach to net savings.[3]

In total, Navigant reviewed the approach to net savings taken by 42 jurisdictions across North America, representing the vast majority of jurisdictions with ratepayer-funded energy efficiency programs. In addition, a review of the approach to net savings in nine selected jurisdictions is discussed in the following section.

[3] Refer to 7.Appendix A for a list of references for methodological resources.

2.3 *Researched NTG Values in Selected Jurisdictions*

To provide the TEC sub-committee with a comprehensive review of researched NTG values Navigant worked with the TEC sub-committee in an iterative process to identify relevant jurisdictions/ programs and accompanying evaluation studies. The research methodology included:

- Review of program evaluations conducted by Navigant and Summit Blue Consulting (acquired by Navigant in 2010),
- Review of program evaluations identified by Navigant staff,
- Review of the Northeast Energy Efficiency Partnerships' Repository of State and Topical EM&V Studies,
- Search of the California Measurement Advisory Council searchable database,
- Search of the Consortium for Energy Efficiency searchable database,
- Review of State and Utility websites for program evaluations and filings,
- General internet searches for program evaluations, and
- Outreach to industry professionals.

This list was revised to develop a shortlist of programs comparable to Union and Enbridge's programs, accounting for factors such as customer segment and program design. Additional studies were excluded due to the methodology employed and/or the applicability of the reported NTG values. [4]

NTG values for programs targeting natural gas savings is the focus of this report due to the greater than expected availability of gas utility studies, as well as combination utility studies where natural gas NTG values were reported separately.

A total of 19 documents[5] were selected covering nine jurisdictions in North America, including: California, Colorado, Massachusetts, Minnesota, New Jersey, New Mexico, Oregon, Washington, and Wisconsin. In some cases, one document reported NTG values for multiple programs, multiple utilities, or multiple program years. In total, 38 distinct NTG values were reported. Table 1 presents the number of distinct values reported across the 19 documents.

[4] Refer to Appendix B for an example of two notable studies/jurisdictions excluded from the analysis.
[5] Refer to Appendix C for an annotated bibliography of these documents.

NAVIGANT

Table 1. Documents Reviewed and Distinct NTG Values Reported

Document Number and Title	Number of Distinct Values Reported	Reason for Including Multiple Values
1. 2004/2005 Statewide Express Efficiency and Upstream HVAC Program Impact Evaluation	4	NTG values reported for 4 utilities: PG&E, SDG&E, SCE, and SCG.
2. 2004-2005 Statewide Nonresidential Standard Performance Contract Program Measurement and Evaluation Study	2	NTG values reported for 2 investor-owned utilities: PG&E and SDG&E.
3. 2006-2008 Retro-Commissioning Impact Evaluation	4	NTG values reported for 4 utilities: PG&E, SDG&E, SCE, and SCG.
4. 2011 Commercial and Industrial Natural Gas Programs Free-Ridership and Spillover Study	6	NTG values reported for 6 utilities: NSTAR, Unitil, New England Gas, National Grid, Columbia Gas, and Berkshire Gas.
5. Evaluation of 2011 DSM Portfolio	2	NTG values reported for 2 programs: Commercial Solutions and SCORE pilot.
6. Fast Feedback Results	3	NTG values reported for 3 programs: Existing Multifamily, Existing Buildings, and Industrial Production Efficiency.
7. Impact and Process Evaluation of the 2006-2007 Building Efficiency Program	2	NTG values reported for 2 program-years: 2006 and 2007.
8. Evaluation of Building Efficiency Program 2004 & 2005	2	NTG values reported for 2 program-years: 2004 and 2005.
9. Impact and Process Evaluation of the 2006-2007 New Building Efficiency Program	2	NTG values reported for 2 program-years: 2006 and 2007.
10. Focus on Energy Evaluation: Business Programs Impact Evaluation Report – Last Quarter of Calendar Year 2009 and First Two Quarters of Calendar Year 2010	2	NTG values reported for 2 program-years: 2009 and 2010.
11. 2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group	1	N/A
12. Evaluation of the Southern California Gas Company 2004-2005 Non-Residential Financial Incentives Program	1	N/A
13. Comprehensive Process and Impact Evaluation of the Business Heating Efficiency Program - Colorado	1	N/A

Document Number and Title	Number of Distinct Values Reported	Reason for Including Multiple Values
14. New Jersey's Clean Energy Program Energy Impact Evaluation: SmartStart Program Impact Evaluation	1	N/A
15. Commercial and Industrial Energy Efficiency Retrofit Custom Programs Portfolio Evaluation	1	N/A
16. Focus on Energy Evaluation: Business Programs – Additional Looks at Attribution	1	N/A
17. Focus on Energy Evaluation: Semiannual Report (Second Half of 2009)	1	N/A
18. Focus on Energy Evaluation: Semiannual Report (First Half of 2009)	1	N/A
19. Achieving Natural Gas Savings Goals: Commercial Heating Programs Heat It Up	1	N/A
Total: 19 Documents Reviewed, 38 Distinct Values Reported		

Source: Navigant analysis.

Navigant reviewed these selected documents to summarize methods used to assess NTG values across these jurisdictions. The following estimates from these studies are reported:

- Net-of-free ridership = 1- FR,
- Net-of-free ridership and participant spillover = 1 – FR + PSO, and
- Net-of-free ridership and all spillover = 1- FR + PSO + NPSO
  (Note: NPSO is non-participant spillover)

This approach conveys information on NTG values based on the common definitions across these studies, and avoids inappropriate comparisons that could result from comparing the studies' reported NTG values when they include different components. Table 2 presents the distribution of the different NTG factors reported across the 38 distinct values.

Table 2. NTG Values Reported

	NTG Values Reported by Adjustment Factor Included	Net-of-NTG Factors
FR	28	38
FR & PSO	3	10
FR, PSO & NPSO	7	7

Source: Navigant analysis.

A total of 28 NTG values reported adjust for FR only, 3 adjust for FR and PSO, and 7 adjust for FR, PSO, and NPSO. The last column shows the information gained from presenting net-of-NTG component values. For example, all 38 of the NTG values reported include values for FR.

Rather than just present the NTG values that adjust for FR only (n=28), the net-of-NTG component values are presented. In this case, (1 – FR) (n=38).[6]

In addition to these studies, Navigant also reviewed the 2008 evaluation of Union and Enbridge's custom projects program conducted by Summit Blue Consulting.[7]

2.4 *Assessing Options for NTG*

Given the uncertainty around NTG values, Navigant applied Decision Analysis methods to illustrate the risks faced by utilities and ratepayers when NTG values are uncertain and provide information on the benefits and costs of choosing one approach to net savings over another.

Navigant took the following steps to conduct the Decision Analysis:

1. Define the benefits of accurate (and costs of inaccurate) NTG values in a general context.
2. Narrow the focus the analysis on the benefits/costs for which Navigant had access to data; specifically, the incentives paid to utilities based on the estimated net savings (m^3) achieved by custom programs.
3. Establish a baseline against which a sensitivity analysis can be conducted where a selected NTG value is assumed to be correct, but in fact is incorrect by some margin of error.[8] The sensitivity analyses were conducted independently by Union and Enbridge and were not verified by Navigant.
4. Conduct a "value of information" analysis by examining the change in incentive payments resulting from better information on NTG values compared to the cost of obtaining the information (e.g., through NTG research).

In addition, Navigant organized the results of the Decision Analysis to provide insight into the tradeoffs from using different approaches to setting an NTG value, ranging from transferring values based on the jurisdictional review to conducting NTG research.

The next section (Section 3) presents an overview of the Union and Enbridge C&I programs to provide context. Following this program overview, Section 4 discusses the regulatory approach and methodological approach to NTG used by different jurisdictions followed by a review of researched NTG values in selected jurisdictions (Section 5). Finally, Section 0 presents the decision analysis for assessing alternate approaches to NTG.

---

[6] Because the documents reviewed contain varying degrees of detail and explanation, the Navigant team applied its best interpretation of these documents to synthesize the available information in a consistent manner.

[7] Summit Blue Consulting. 2008. *Custom Projects Attribution Study.* Union Gas Limited and Enbridge Gas Distribution, October 27, 2008.

[8] These first three steps are part of a "loss function" analysis which identifies the costs of selecting one NTG value when another value is the actual value.

3. Overview of Union and Enbridge Custom Programs

Union and Enbridge have been delivering natural gas DSM programs for over 10 years, including custom programs for the C&I sectors. This section provides an overview of these programs.

3.1 Union Custom Programs

Union offers the Custom Savings Program to C&I customers. Within the custom program umbrella there are numerous program offerings providing a combination of technical assistance and financial incentives:

- **Engineering Feasibility Study.** These comprehensive engineering analyses and assessments include both whole facility and end-use focused studies. Example projects include thermal surveys, HVAC audits, energy audits, and energy benchmarking.
- **Steam Trap Survey.** These studies focus exclusively on the use and efficiency of steam traps, and seek efficiencies in the discharge of condensation, air, and other non-condensable gases without losing steam.
- **Process Improvement Study.** This offering targets industrial facilities through comprehensive process improvement studies conducted by industry-specific production and energy utilization experts. Example projects include steam plant audits, process integration analyses, heating integration studies, and process operation improvement studies.
- **Integrated Energy Management Systems.** This program offering provides technical assistance and financial incentives to industrial customers for the installation of an integrated management system.
- **Customer Education.** This program provides education, training, and technical assistance to C&I customers.
- **New Equipment.** Technical assistance and financial incentives are provided to C&I customers to support the installation of new energy efficient equipment and processes. Examples of measures include furnaces, HVAC, heat recovery, controls, insulation, and building envelope.
- **Runsmart Building Optimization.** Technical assistance and financial incentives are provided to commercial customers (e.g., education, healthcare, offices, multi-unit residential, and entertainment) for building optimization. Examples of projects include verifying dampers and valves on air handling units, calibrating sensors and instrumentation, and insulation.

- **Operation and Maintenance.** This program offering provides technical assistance and financial incentives to C&I customers for operation and maintenance of existing measures. Typical projects include repairs to HVAC systems, hot water systems, insulation repairs, and steam system repairs.
- **Boiler Tune-Up.** Technical assistance and financial incentives are provided to industrial customers for a boiler tune-up. Boilers must have output of less than 25,000 pounds per hour or 800 BHP.
- **Meters.** Technical assistance and financial incentives are provided to industrial customers for the installation of natural gas, steam, or hot-water meters.
- **Infrared Anti-Condensate Plastic.** This program offering provides technical assistance and financial incentives to industrial customers for the installation of infrared anti-condensate plastic for a greenhouse.
- **Demonstration of New Technologies.** Technical assistance and financial incentives are provided to C&I customers for adopting new technologies that save natural gas.

3.2 *Enbridge Custom Programs*

Enbridge offers two custom C&I programs:

- **Commercial Custom Savings Program** provides both technical assistance and financial incentives to medium to large-sized new and existing commercial customers for energy efficient custom gas projects. Examples of custom measures include boilers, building automation systems, variable frequency drives, and demand control ventilation.
    1. The *Existing Buildings* program offering primarily focuses on projects with multiple technologies and requires technical assistance throughout the development of the project.
    2. Two new initiatives, launched in 2012, (*Energy Compass and Run It Right*) encourage a continuous improvement strategy for large commercial customers. These program offerings provide technical assistance by offering an energy efficiency diagnostic service and assisting with the implementation of low and no-cost operational improvements.
- **Industrial Continuous Energy Improvement Program** aims to reduce the natural gas use of medium to large-sized industrial customers through a continuous improvement approach. This approach includes five steps, providing both technical assistance and financial incentives for the implementation of energy efficiency projects:
    1. *Knowledge Development* involves educating customers through workshops and publications.
    2. *Opportunity Identification* involves providing technical assistance to customers in identifying energy efficiency opportunities.

3. *Measurement* provides technical assistance to identify and measure the information needed to make a decision regarding energy efficiency opportunities. Financial incentives are available for measurement equipment.
4. *Engineering Analysis* provides technical assistance to customers in quantifying the benefits and costs associated with an energy efficiency opportunity. Financial incentives are available if a third party consultation is required.
5. *Action and Implementation* provides technical assistance and financial incentives for energy efficiency projects.

Examples of projects include industrial process heat systems, steam systems, and heating and ventilation.

4. Approach to Net Savings

This section presents the findings from the jurisdictional review of the approach taken to net savings, as well as the availability of performance incentives. This section begins with a review of 42 jurisdictions in the United States and Canada, representing the vast majority of jurisdictions with ratepayer-funded energy efficiency programs. This is followed by a closer look at the nine jurisdictions selected for further review. The final section summarizes the findings that are most relevant to Union and Enbridge.

4.1 Jurisdictional Review

Table 3 presents a summary of the approach to net savings used in the 42 jurisdictions, including the treatment of a FR adjustment and whether spillover is considered.[9] The table also presents information on whether jurisdictions offer utility performance incentives for meeting their savings goals, though, as indicated below, these goals are linked to either *gross* or *net savings*. Following is a summary of key findings:

- One-third (33%) of the jurisdictions reviewed **do not adjust gross savings** for either FR or spillover; however, some of those states may conduct some NTG research to inform future program design. Half of the U.S. states that do not adjust gross savings provide performance incentives for utilities to achieve energy efficiency program goals or have a performance incentive pending.

- Relatively few (14%) of the jurisdictions reviewed use a **deemed approach** to NTG; the deemed NTG values may be determined at a portfolio level (ranging from 0.7 to 0.9) or on a measure-by-measure basis (as in California, Vermont, and Nevada). These deemed NTG values are typically developed after NTG research has been conducted through program impact evaluations, and are revised on a regular basis through negotiations between utilities and regulators (often informed by additional NTG research). Over three-quarters (83%) of the U.S. states that use a deemed NTG approach provide performance incentives for utilities to achieve energy efficiency program goals.

- Nearly half of all jurisdictions reviewed take a **research-based approach** to NTG analysis. The vast majority of those jurisdictions consider spillover in some capacity, at least for some program types, though spillover is still quantified much less often than FR. Both FR and spillover are most commonly estimated through a self-report (participant survey) approach, though econometric methods (e.g., billing analysis) and market share modeling approaches are occasionally used. Nearly three-quarters of the U.S. states that take a research-based NTG approach provide performance incentives for

[9] Note that within a given jurisdiction, the treatment of spillover may vary by program type (including whether participant, non-participant, or both types of spillover is researched), and evaluators may investigate the possibility of spillover but find that no spillover is occurring or that it cannot be quantified with enough precision to obtain regulatory approval. Thus, this column reflects jurisdictions which consider the possibility of spillover but have not necessarily quantified and received regulatory approval for spillover savings estimates.

utilities to achieve energy efficiency program goals or have a performance incentive pending.

Table 3. NTG Approaches, Treatment of Free Ridership and Spillover, and Availability of Performance Incentives by Jurisdiction

Jurisdiction	NTG Approach*	Free-Ridership Adjustment	Spillover Considered?	Performance Incentives?	Notes
Hawaii	Deemed (0.7)			Yes	
Arkansas	Deemed (0.8)			Yes	
Michigan	Deemed (0.9)			Yes	Some NTG research conducted but not currently required by regulators.
California	Deemed (varies by measure, 0.5 for custom gas measures)			Yes	Research conducted to inform deemed NTG values.
Nevada	Deemed (varies by measure)				Some NTG research conducted.
Vermont	Deemed (varies by measure)			Yes	
British Columbia	Researched	Yes	Yes		Deemed NTG of 1.0 used until researched.
Nova Scotia	Researched	Yes	Yes		
Colorado	Researched	Yes	Yes	Yes	
Connecticut	Researched	Yes	Yes	Yes	Gross savings are used to evaluate whether goals have been met.
Florida	Researched	Yes	Yes	Pending	
Georgia	Researched	Yes	Yes	Yes	
Illinois	Researched	Yes	Yes		
Indiana	Researched	Yes	Yes	Yes	
Kansas	Researched	Yes		Pending	
Maine	Researched	Yes	Yes		
Massachusetts	Researched	Yes	Yes	Yes	
Missouri	Researched	Yes	Yes	Pending	
New Hampshire	Researched		Yes	Yes	
New Mexico	Researched	Yes		Yes	

Jurisdiction	NTG Approach*	Free-Ridership Adjustment	Spillover Considered?	Performance Incentives?	Notes
New York	Researched	Yes	Yes	Yes	Deemed NTG of 0.9 used for programs without recent evaluations.
Oregon	Researched	Yes	Yes		
Pennsylvania	Researched	Yes	Yes		Gross savings are used to evaluate whether goals have been met.
Rhode Island	Researched		Yes	Yes	
Utah	Researched	Yes	Yes	Pending	
Wisconsin	Researched	Yes	Yes	Yes	
Wyoming	Researched	Yes	Yes		
Arizona	No NTG adjustment			Yes	
Delaware	No NTG adjustment				
District of Columbia	No NTG adjustment				
Idaho	No NTG adjustment			Pending	Some NTG research conducted but not required by regulators.
Iowa	No NTG adjustment				
Kentucky	No NTG adjustment			Yes	
Maryland	No NTG adjustment				
Minnesota	No NTG adjustment			Yes	
Nebraska	No NTG adjustment				
New Jersey	No NTG adjustment				
North Carolina	No NTG adjustment			Yes	
Ohio	No NTG adjustment			Yes	
Texas	No NTG adjustment			Yes	

Jurisdiction	NTG Approach*	Free-Ridership Adjustment	Spillover Considered?	Performance Incentives?	Notes
Washington	No NTG adjustment				Some NTG research conducted but not required by regulators.
South Dakota	Varies by utility	Yes	Yes		

* Deemed NTG values are pre-determined values typically developed after NTG research has been conducted through program impact evaluations. Researched NG values are most commonly estimated through a self-report (participant survey) approach, though econometric methods (e.g., billing analysis) and market share modeling approaches are occasionally used. *Source:* Navigant analysis of various resources including utility websites, regulatory agency websites, websites of research/advocacy groups, and studies that previously surveyed the approach to net savings (Appendix A).

4.2 *Selected Jurisdictions*

As noted in the Methodology section, Navigant reviewed a total of 19 documents that researched NTG. These documents represent nine jurisdictions, including: California, Colorado, Massachusetts, Minnesota, New Jersey, New Mexico, Oregon, Washington, and Wisconsin.

While documents that research NTG were identified, the approach to net savings in these selected jurisdictions varies as shown in Table 4. Most notably, three of the jurisdictions make no NTG adjustment and one jurisdiction deems NTG even though NTG research is being conducted. Also note that three of the nine jurisdictions do not have performance incentives.

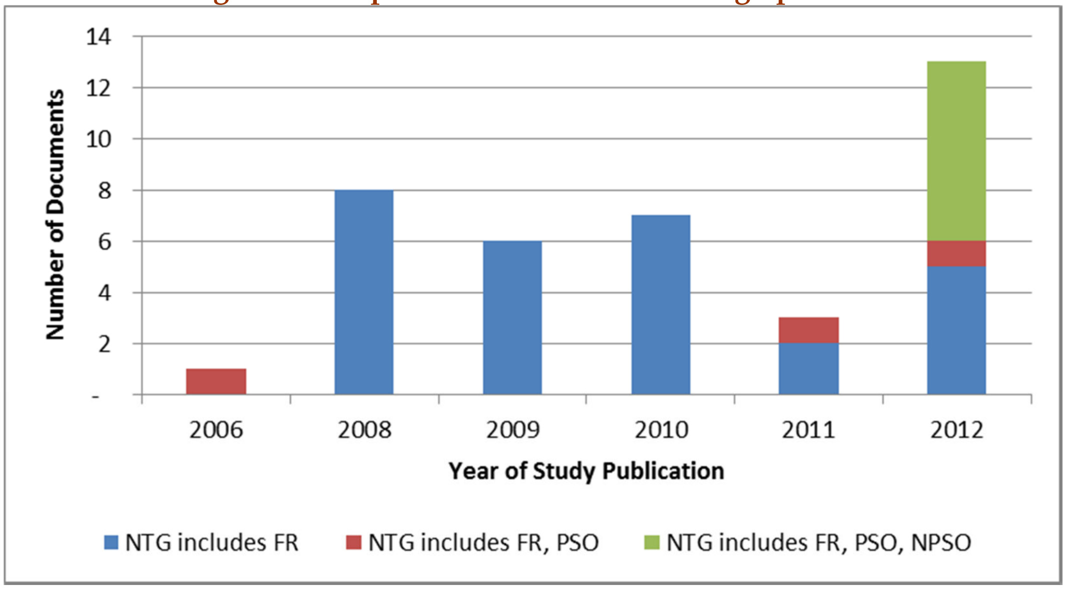
Table 4 . Approach to Net Savings in Selected Jurisdictions

Deemed	Researched Adjusts for Free Ridership and Spillover is Considered	No NTG Adjustment
California (0.5 for custom gas measures)	Colorado, Massachusetts, New Mexico (FR only), *Oregon*, and Wisconsin	Minnesota, *New Jersey*, and *Washington*

**Italics* indicate that the jurisdiction does not have performance incentives. *Source:* Navigant analysis.

Regional or temporal trends in whether participant and NPSO were also considered. Figure 1 presents the number of studies that include free-ridership, PSO, and NPSO by the year of study publication. Based on the sample of studies conducted in the selected jurisdictions, there is a clear trend towards including participant and NPSO in calculating NTG in recent years.

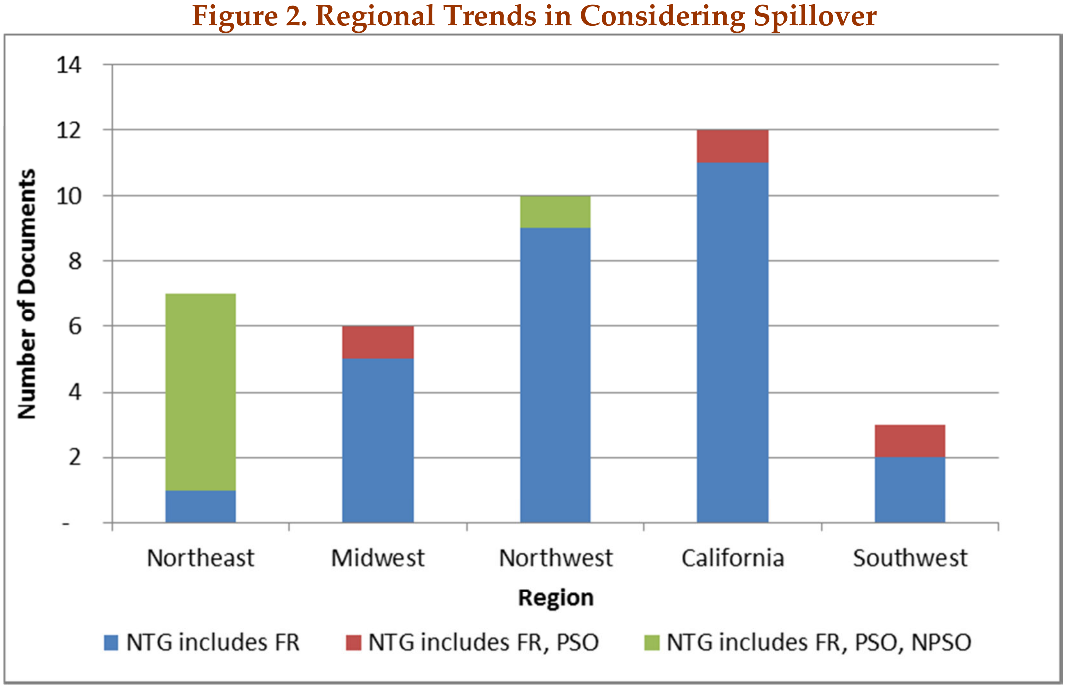
Figure 1. Temporal Trends in Considering Spillover



Source: Navigant analysis.

Figure 2 presents the number of studies that include free-ridership, PSO, and NPSO by region of the United States. Based on the sample of studies conducted in the selected jurisdictions, it appears that all regions consider PSO in calculating NTG values.

Figure 2. Regional Trends in Considering Spillover



Source: Navigant analysis.

4.3 Application to Union and Enbridge

Based on the jurisdictional review nearly half of the jurisdictions with rate-payer funded energy efficiency program conduct NTG research. Among the 33% that do not adjust gross savings some research is being conducted. For example, three of the nine jurisdictions selected for further review do not adjust gross savings while another one deems – yet NTG research is being conducted.

Trends in the included NTG factors are also identified. Among the nine selected jurisdictions there is a clear trend towards including both participant and NPSO in recent years, and that it is not a regional phenomenon. The next section of this report summarizes the researched NTG values resulting from the review of research conducted in the nine selected jurisdictions.

5. Researched NTG Values in Selected Jurisdictions

In this section Navigant summarizes the 38 NTG values reviewed in the nine selected jurisdictions. As described in Section 2.3, the NTG values presented are net-of-NTG factors. All values represent gas values, unless specified otherwise.

A summary of the studies' findings across the following categories are presented:

- First, a high level summary of the NTG values for non-residential natural gas programs is provided. To provide context for these values we examine how these values vary with the document number, region, program year, and the analytic rigor of the methodology used. We also provide a comparison of the natural gas NTG values to the electric NTG values reported in the same documents.

> ***Definitions***
>
> NTG values presented in this section represent "Net-of-NTG Factors."
>
> - NTG value including free ridership, NTG = (1-FR),
> - NTG value including free ridership and participant spillover, NTG = (1-FR+PSO), or
> - NTG value including free ridership and spillover, NTG = (1-FR+PSO+NSPO), where NPSO represents non-participant spillover.

- Next, the NTG values based on a variety of program characteristics, including program type, customer segment, utility-type, region, approach to program marketing, and program maturity are summarized.[10]
- The final section summarizes the findings that are most relevant to Union and Enbridge.

It is important to keep in mind that the NTG values presented in this section are the result of research conducted for different programs, in different program environments, and using different methodologies. As a result, interpretation of trends should be made with caution - differences in NTG values may reflect true differences in FR and spillover, or may simply reflect differences in evaluation methodologies, even among similar programs (Saxonis 2007).

5.1 *Summary of NTG Values*

Figure 3 summarizes net of NTG component values.[11] Some key patterns are evident in this Figure:

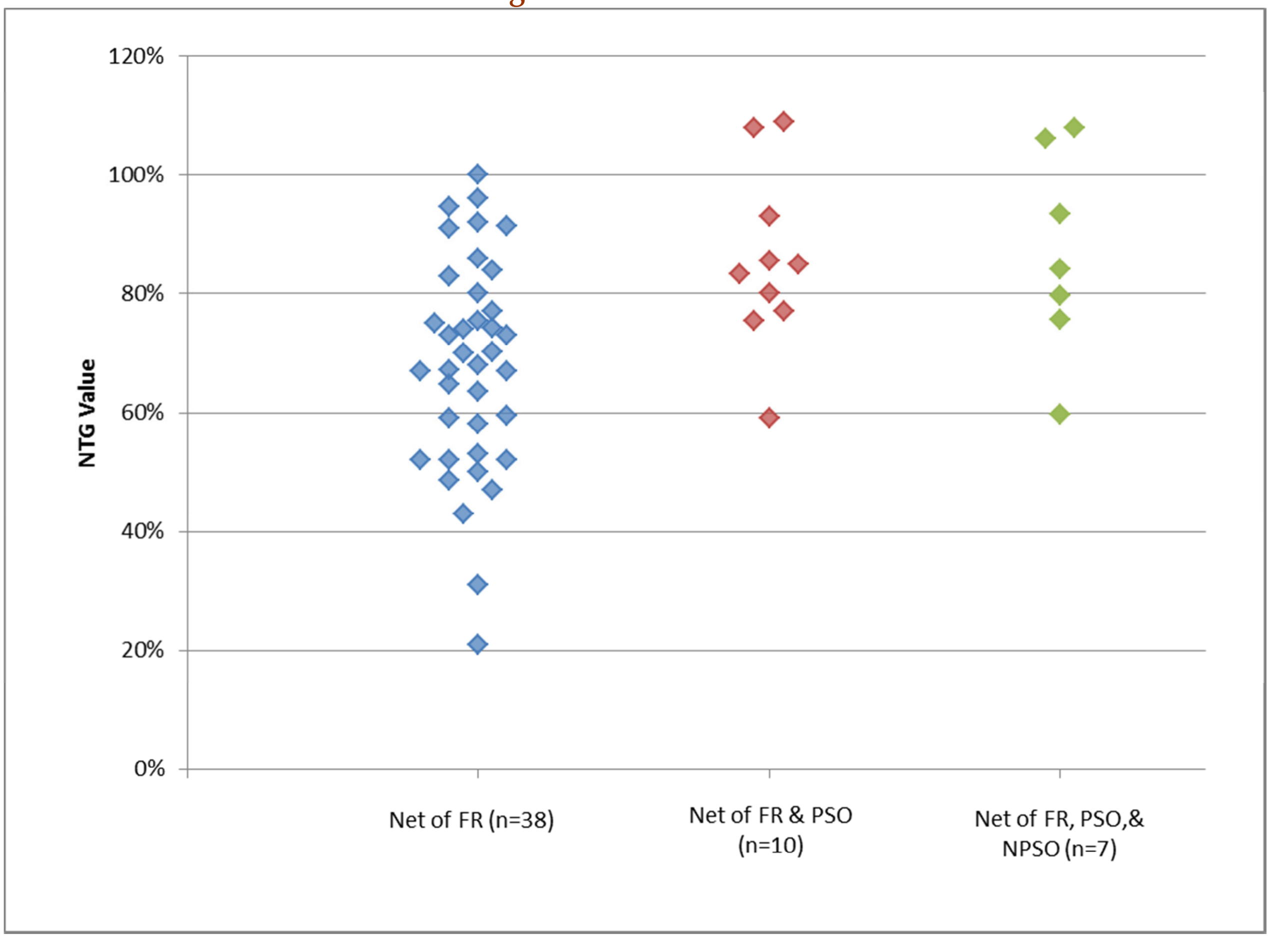
---

[10] Summarizing NTG values by various categories limits the sample sizes. As a result, caution should be used in interpreting NTG values.

[11] By presenting net-of-NTG component values, a distinct result reported in a document may be represented by multiple data points in the figures below. For example, if free ridership, PSO, and NPSO are considered, three data points will appear in the figure: the net-of-FR value, the net-of-FR & PSO value, and the net-of-FR, PSO & NPSO value.

- While the dispersion of net-of-free ridership values is quite large, ranging from 21% to 100%, the majority of values appear to "cluster" between 40% and 90%.
- There are only a few studies at the extremes of the range of net-of-free ridership values. One result reports high levels of free-ridership (79%) with another reporting zero free-ridership.[12]
- The average net-of-free ridership value is 68%.
- As expected, NTG values are larger when considering spillover. Average net-of-free ridership & PSO value is 86% and average net-of-free ridership & spillover value is 87%, suggesting that NPSO is small for non-residential gas programs.[13]

Figure 3. NTG Values



Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG values (program-utility-year combinations) reported in the 19 studies.

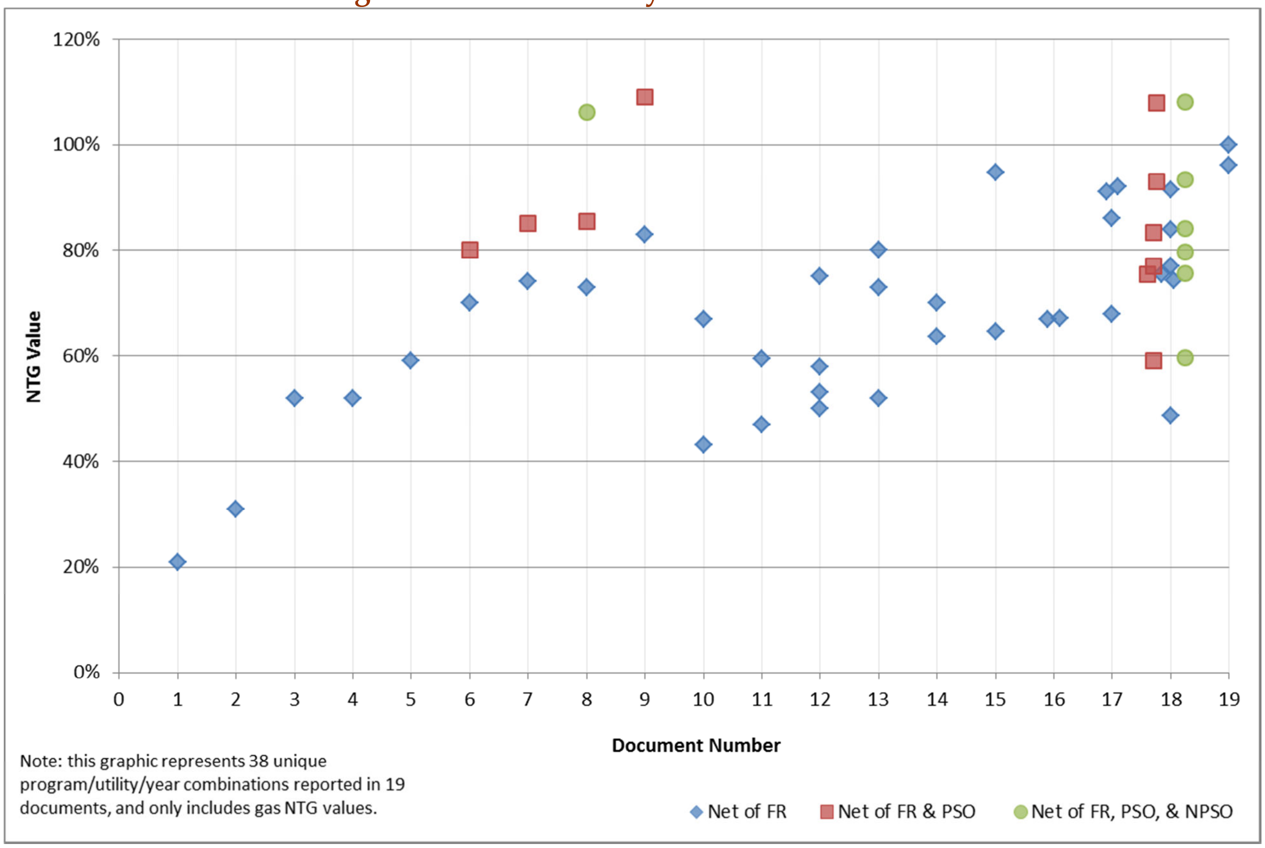
[12] Zero free-ridership was reported for a small pilot program (n=30) offering custom and prescriptive incentives targeted at K-12 school districts. 79% free-ridership was reported for a retrofit program in its third program year. The sample size (n=18) represents 75% of participants with natural gas measures and 10% of total program participants. Both studies relied on self-report methods.

[13] 5 of the 7 data points for NPSO report values of less than 1% with another reporting 2.6% (all values reported by the same study). The remaining data point reports NPSO of 21% with a corresponding PSO value of 13%).

To further examine trends in NTG values, Figure 4 summarizes the distinct NTG values reported by each document. There are two key findings:

- Only two documents report net-of-FR values below 40%.
- Net-of-FR values that exceed 90% are reported by just four documents and generally exhibit a clustering of multiple values. For example, document number 19 reports two distinct NTG values, both of which are larger than 90%.

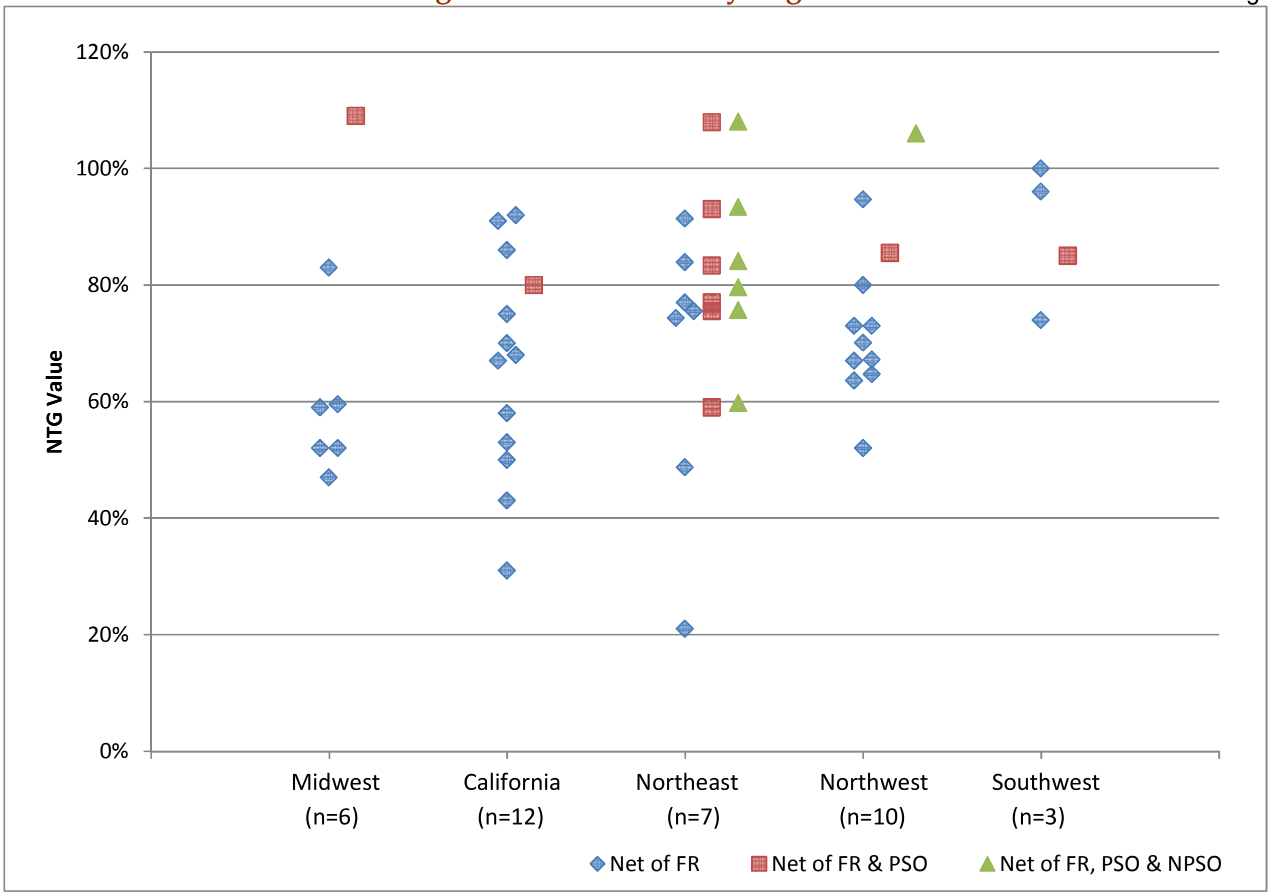
Figure 4. NTG Values by Document Number



Source: Navigant analysis.

Figure 5 summarizes NTG values by region. No clear regional trends emerge except it appears there is a clustering of net-of-FR values in the Northwest around 70%. These values represent evaluations of multiple program-years of two programs, with evaluations conducted by multiple evaluators.

Figure 5. NTG Values by Region

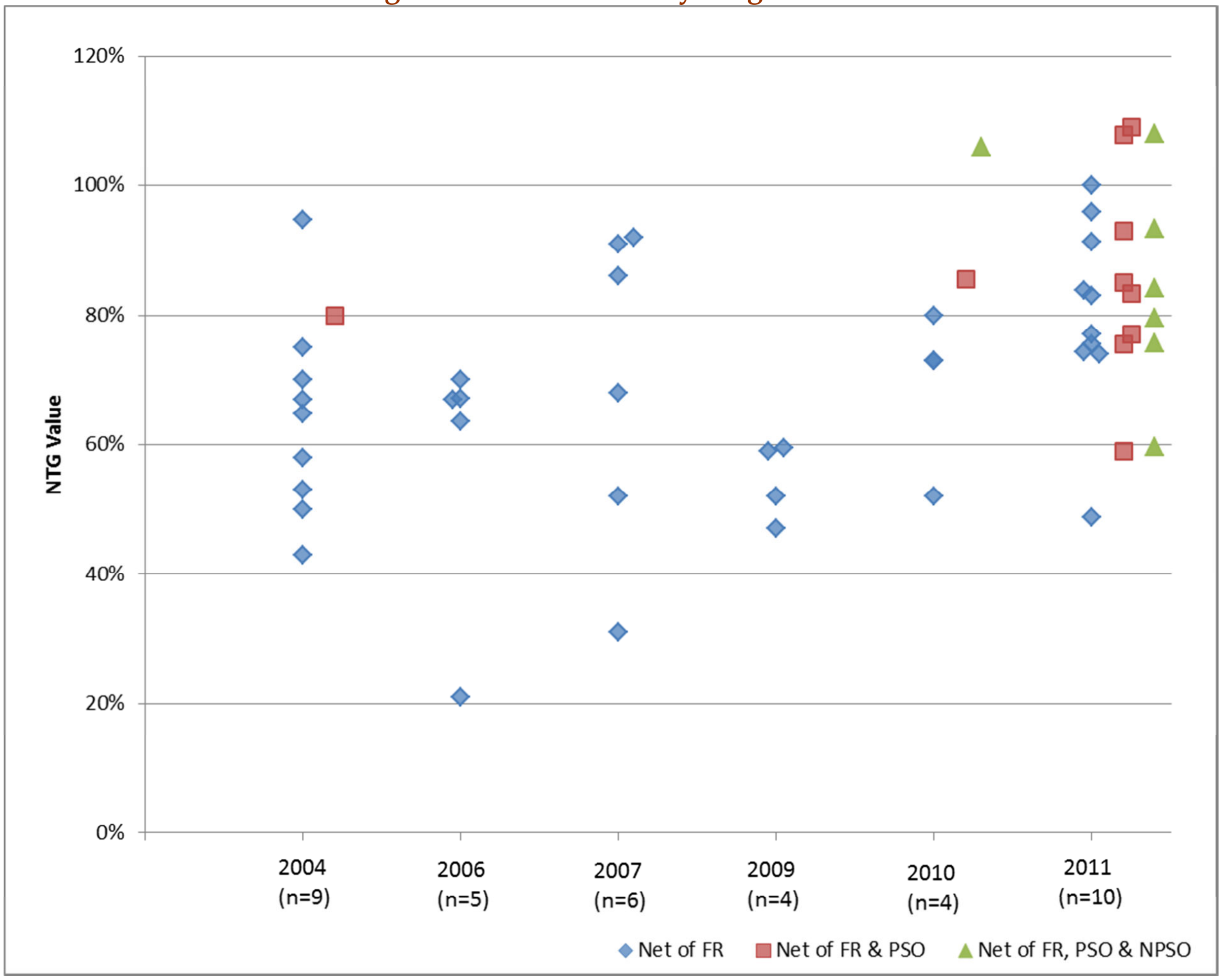


Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) in each region; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

Economic conditions may influence NTG values though few longitudinal studies have been conducted to reveal with certainty how FR and spillover are influenced. Saxonis (2007) identifies research conducted in the 1990's that suggest FR is lower during economic downturns. To ensure that trends in NTG values are not driven by specific economic conditions, Navigant explored whether NTG values vary by program year in Figure 6.[14] While there is a slight upward trend in the net-of-FR estimates, it is not large enough to cause concern about using average values if the TEC decides to do so.

[14] When two program years were evaluated, the first program year is used. For example, if a study evaluates program years 2004-2005, the NTG value is recorded for 2004.When three program years were evaluated, the middle program year is used. For example, if a study evaluates program years 2006-2008, the NTG value is recorded for 2007.

Figure 6 . NTG Values by Program Year



Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) by program year; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

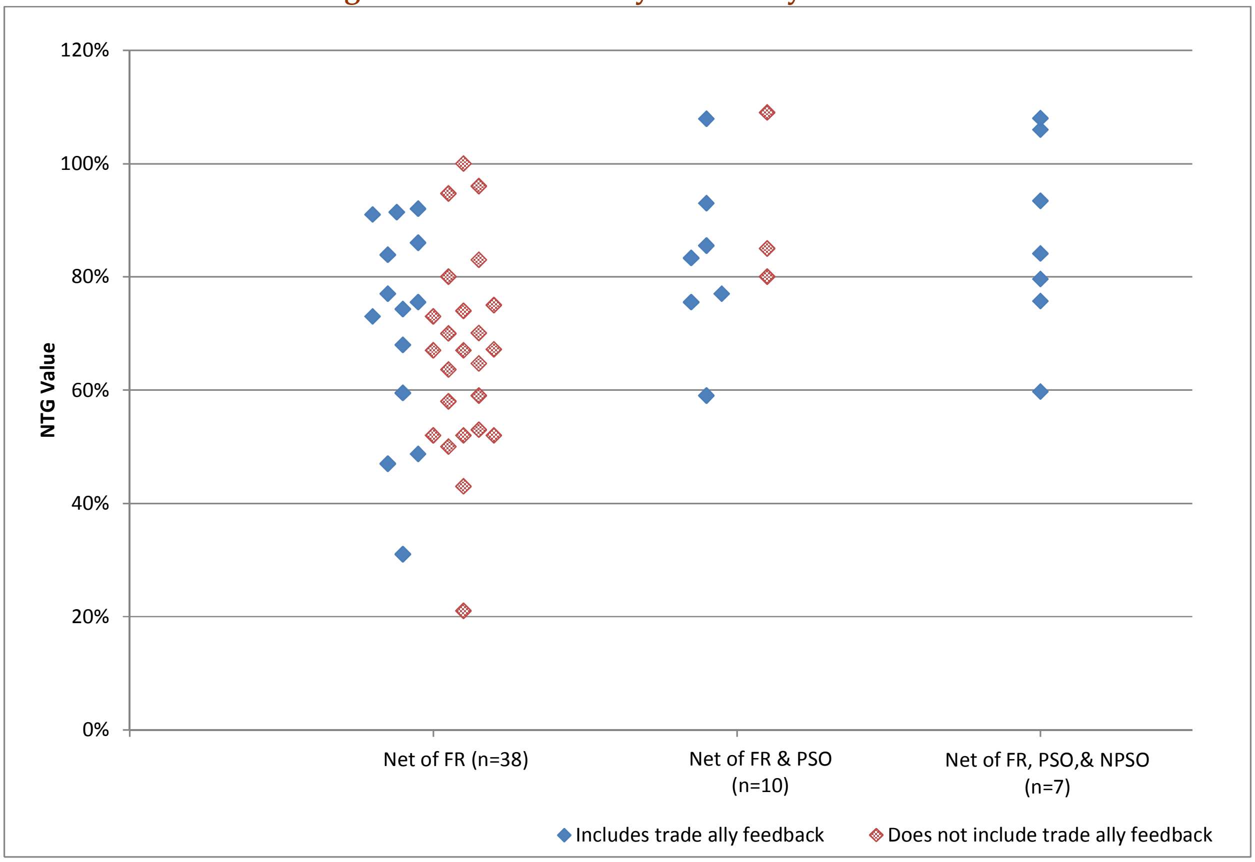
To provide further context to this summary of NTG values Navigant explored whether there are trends in NTG values based on the analytic rigor of the methodology, but were limited in our efforts due to a lack of data. For example, the sample size for most of the results was identified, but the documents did not report population size or the fraction of energy savings that the sample size represents. Without context for the sample size, information on how NTG values vary with sample size provides little insight.[15]

Instead, Navigant uses a proxy for the analytic rigor of the methodology based on data that is available, namely, whether the evaluators used enhanced self-report methods in the form of trade ally feedback. Figure 7 summarizes NTG values differentiating between whether trade ally feedback was incorporated in the NTG calculation. Net-of-free ridership values appear to

[15] Refer to Appendix D for information on sample size.

cluster at slightly larger values when incorporating trade ally feedback. This is not unexpected as trade ally feedback often decreases FR because trade allies have more insight about the full extent of the program's influence on the market.

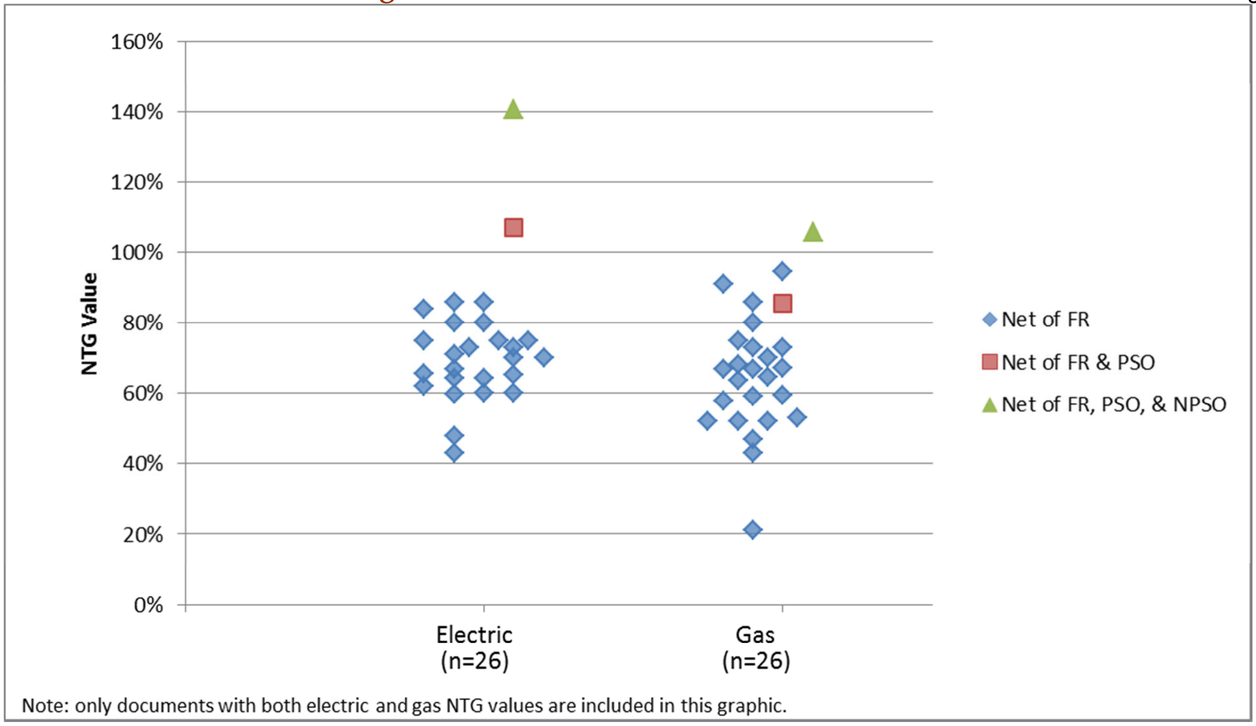
Figure 7. NTG Values by Trade Ally Feedback



Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG values (program-utility-year combinations) reported in the 19 studies.

Comparing gas NTG values to electric NTG values may also provide additional insight. Many of the documents reviewed target both electric and gas measures, but report NTG values for electric and gas measures separately. Figure 8 compares electric NTG values to gas NTG values for those documents that report both electric and gas NTG values. Net of FR values appear to cluster for both gas and electric, but the clustering of gas values is slightly wider than electric. Average net-of-free ridership values are similar, 69% for electric and 65% for gas.

Figure 8. Electric versus Gas NTG Values



Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) for each fuel type; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

The following section examines whether NTG values vary by features of program design and delivery.

5.2 *Summary Based on Program Characteristics*

In this section, Navigant summarizes NTG values based on various characteristics of program design and delivery. In particular, variation in NTG values is examined based on:[16]

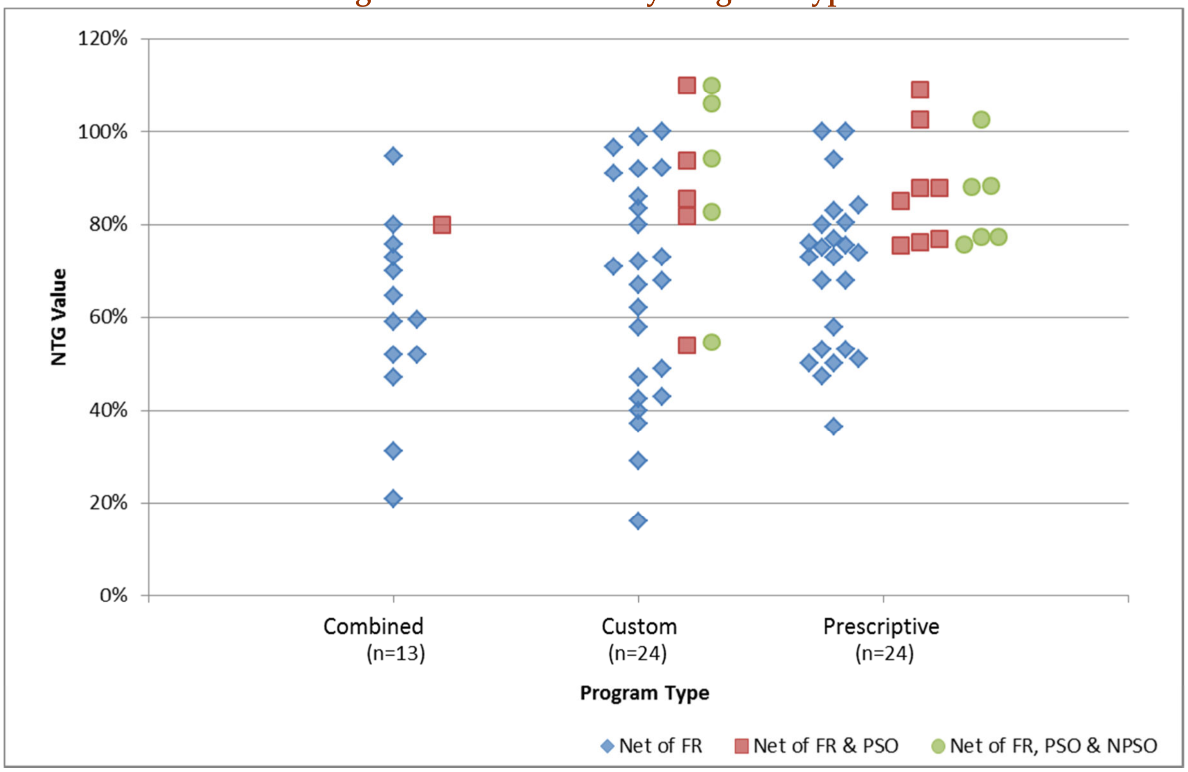
1. **Program-type**, differentiating between custom, prescriptive, and both.
2. **Customer segment**, differentiating between commercial, industrial, agricultural, institutional, and multi-sector.
3. **Utility-type,** differentiating between utilities/organizations that offer electric and gas versus those that offer gas-only.
4. **Program maturity,** differentiating by the number of years since program inception.

[16] Navigant explored other characteristics of program design, such as incentives as a percent of incremental cost, extent of design assistance throughout the program, program objectives, and more, however, because most studies did not provide this level of detail on the programs they were not included in the analysis.

5. **Program marketing strategy**, differentiating between a direct marketing/outreach, channel/partners, and both.

Figure 9 summarizes NTG values by program type (custom, prescriptive, or both).[17] Custom net-of-FR values exhibit a wider dispersion relative to prescriptive values. Excluding some outlier custom values, the ranges are fairly similar but the prescriptive values exhibit more clustering between 50% and 85%, whereas custom values do not appear to cluster in any particular range of values.

Figure 9. NTG Values by Program Type



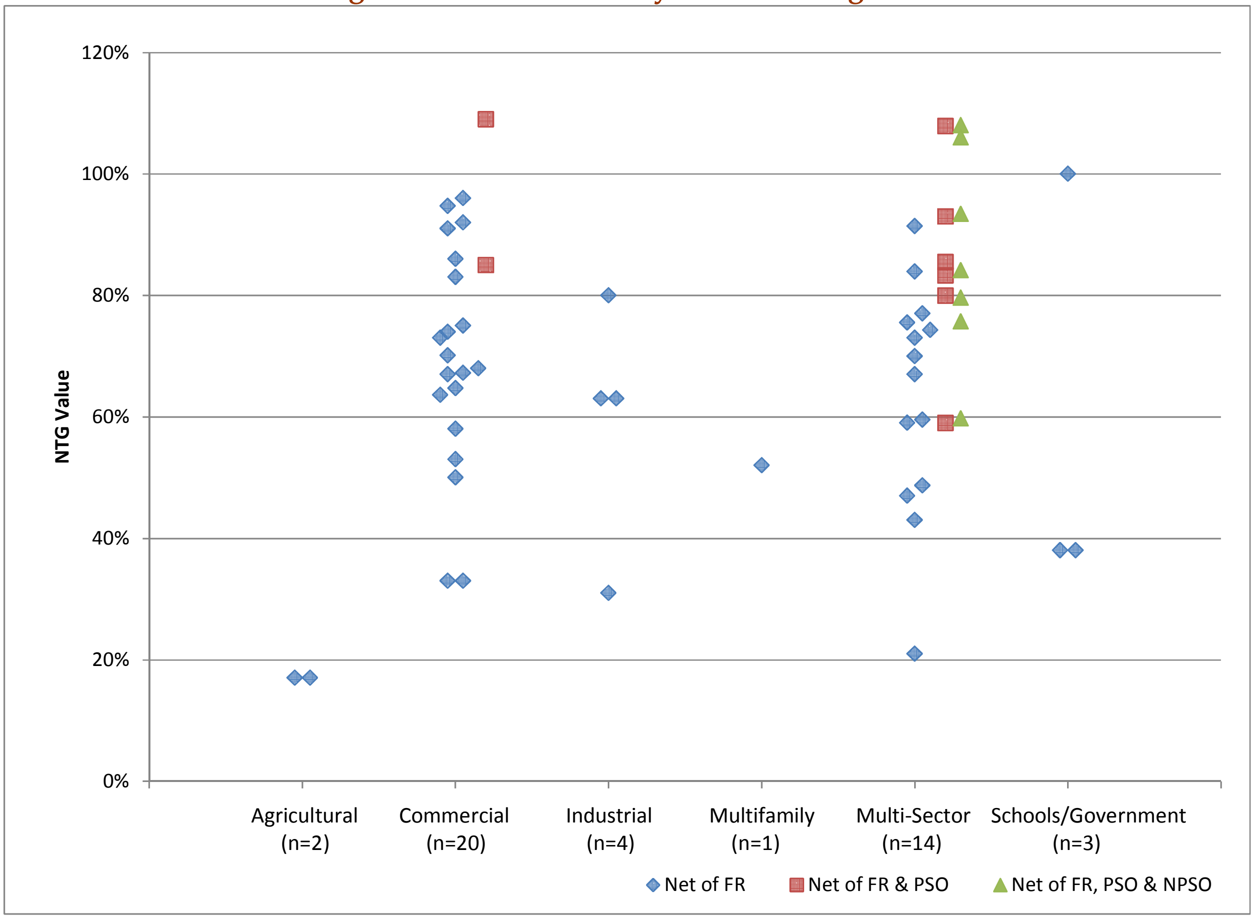
Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) for each program type; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

Figure 10 summarizes NTG values by customer segment.[18] Most of the programs included in this review are targeted at the commercial sector or are classified as multi-sector programs. While there is a wide dispersion of NTG values, the majority of values are found within the 60% and 80% range.

[17] In an effort to identify whether there are trends in NTG values by program type, when a NTG value was disaggregated into custom and prescriptive categories, these NTG values were included separately, resulting in a total of 61 data points for this analysis.

[18] In an effort to identify whether there are trends in NTG values by customer segment, when a NTG value was disaggregated into customer segments, these NTG values were included separately, resulting in a total of 44 data points for this analysis.

Figure 10. NTG Values by Customer Segment

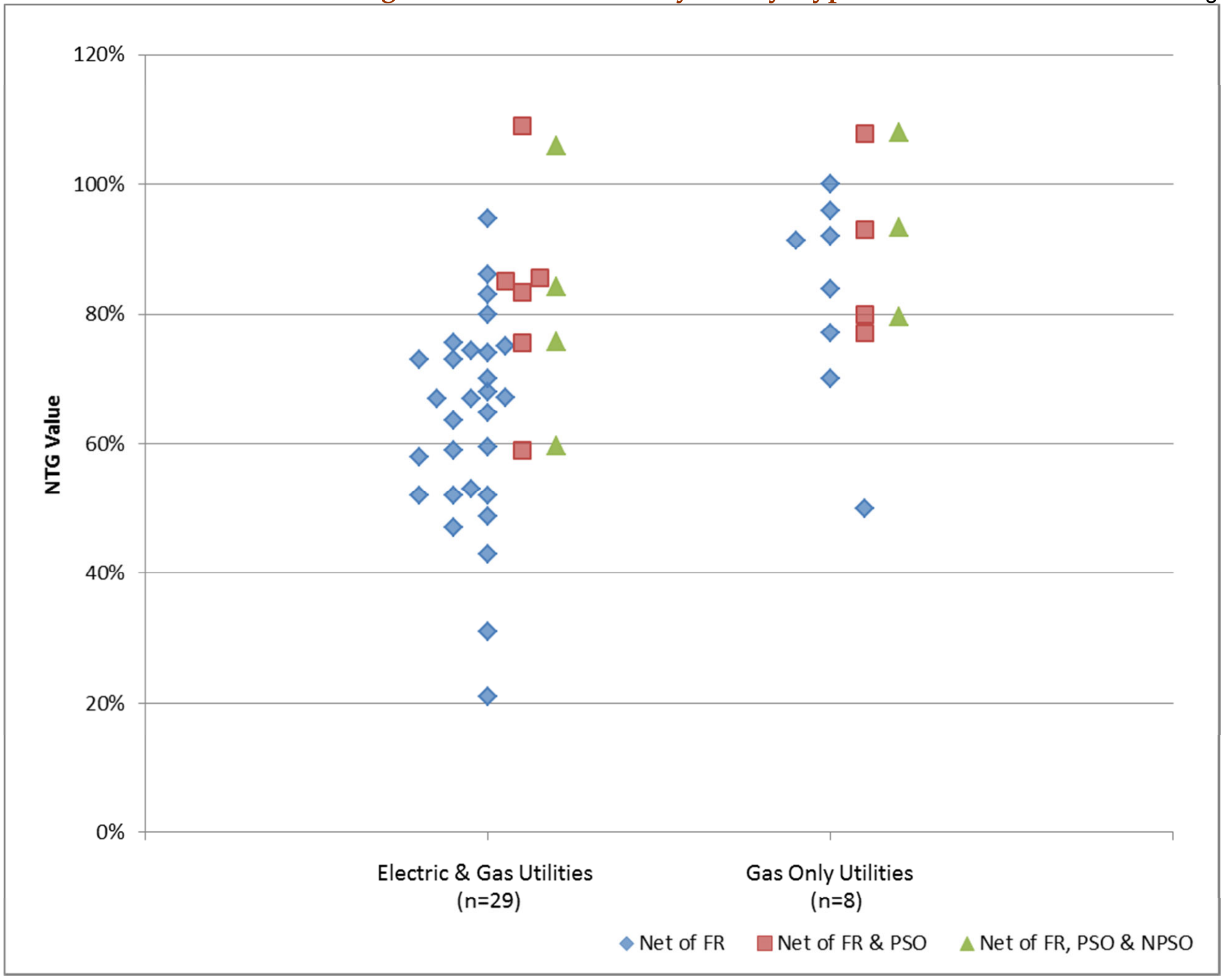


Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) for each segment; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

Figure 11 summarizes NTG values by utility-type (e.g., gas only, electric and gas).[19] Of the documents reviewed, more programs are offered by electric and gas utilities relative to gas-only. With only a few distinct net-of-FR values for gas-only utilities, comparisons across utility-types should be made with caution. Nevertheless, there appears to be a trend of lower FR and higher NTG values for programs offered by gas-only utilities.

[19] Note that the values presented are gas NTG values.

Figure 11. NTG Values by Utility-Type

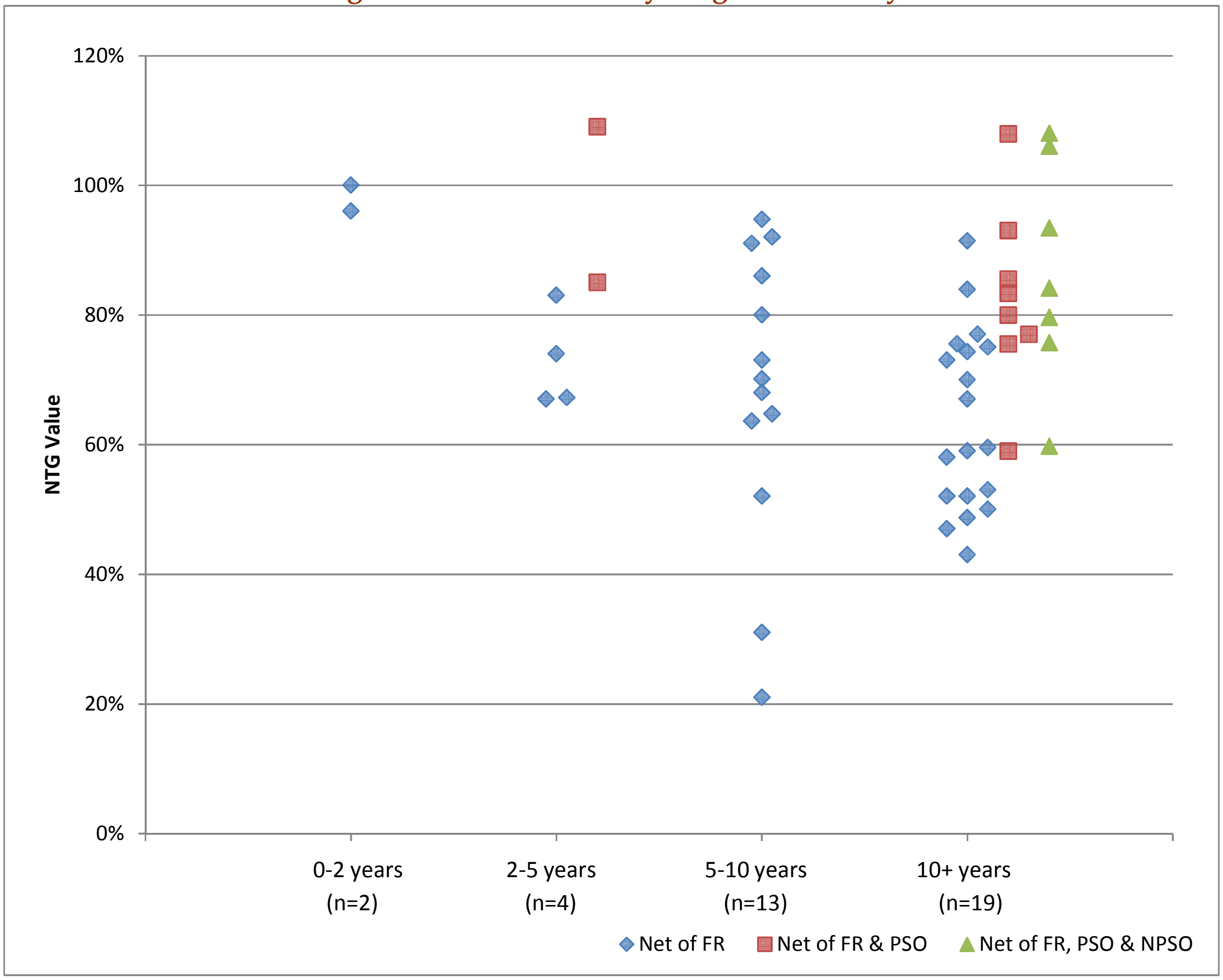


Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) for each utility-type; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable). Total sample size is 37 instead of 38 because one utility is electric only but reported NTG values for gas savings from electric programs, specifically a retrofit program.

Navigant also explored whether NTG values varied with program maturity and program marketing strategy. Figure 12 summarizes NTG values by program maturity. The majority of programs are in at least their fifth program year, and while the sample size of programs with less than 5 years' experience is limited, there appears to be a trend of lower NTG values (and higher FR) as program experience increases. This finding is not unexpected as markets transform over time raising awareness and knowledge of the benefits of energy efficiency among potential resulting in higher degrees of FR. Jurisdictions which only adjust for FR can be especially prone to declining NTG values over time because what appears like FR in a program's later years may actually be evidence of spillover or market transformation from the program's earlier market interventions.

NAVIGANT

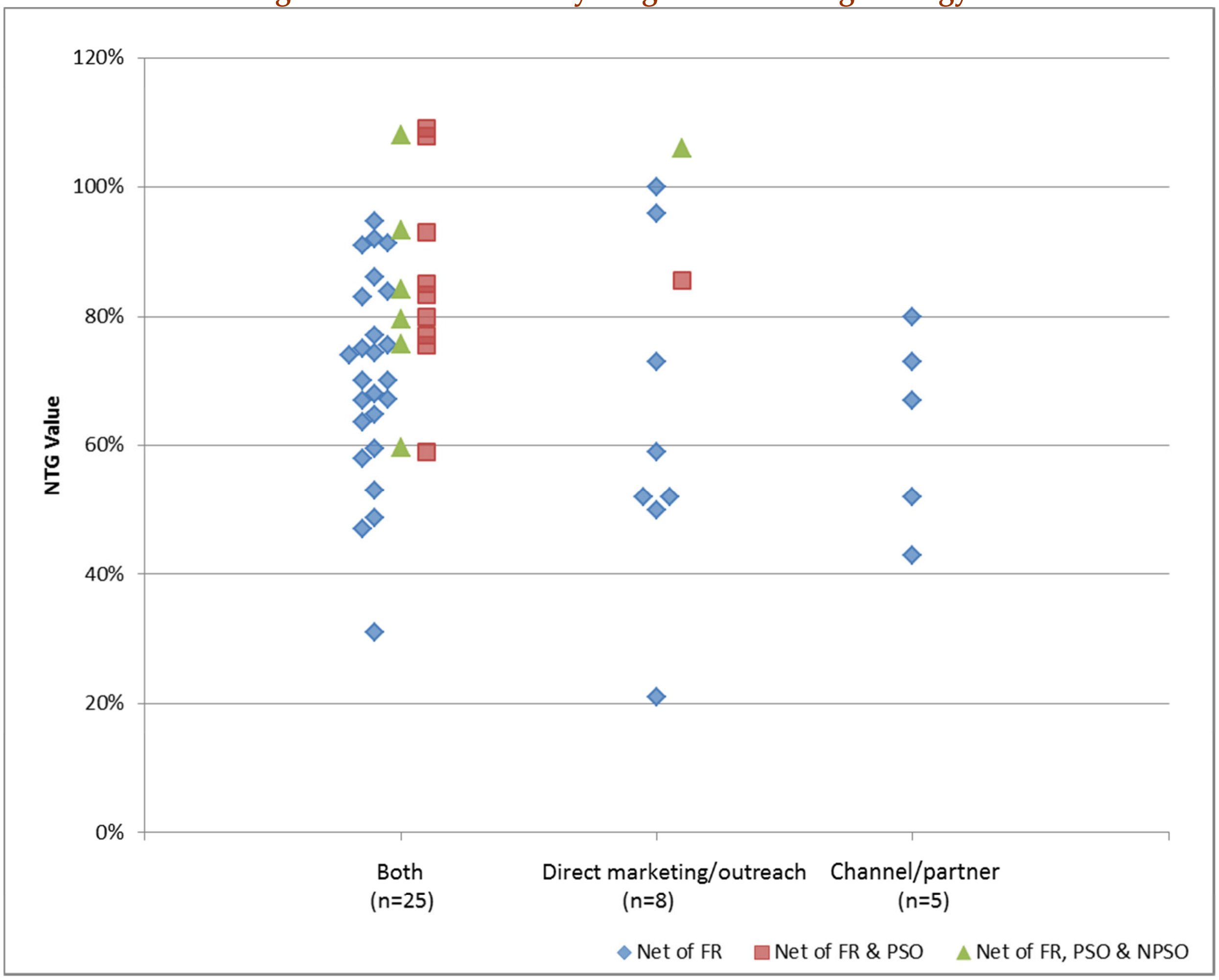
Figure 12. NTG Values by Program Maturity



Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) by program maturity; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

Figure 13 summarizes NTG values by program marketing strategy. The majority of programs adopted both a direct marketing/outreach strategy and a channel/partner strategy. As a result, the distribution of NTG values is similar to the high-level summary depicted in Figure 3. Note that the extreme net-of-FR values of 100% and 21% are for programs with a direct marketing/outreach strategy.

Figure 13. NTG Values by Program Marketing Strategy



Source: Navigant analysis. Note that the sample size (n) represent the number of unique NTG values (program-utility-year combinations) by program marketing strategy; the number of data points in the figure exceed the sample sizes because NTG findings are presented as net-of-free ridership, net-of-free ridership & PSO (if applicable), *and* net-of-free ridership, PSO & NPSO (if applicable).

5.3 *Application to Union and Enbridge*

In 2007-2008 Navigant (formerly Summit Blue Consulting) conducted the first attribution study of Union and Enbridge's custom C&I programs to evaluate FR and spillover effects. Table 5 presents the NTG values as well as the values of the individual NTG components.[20]

[20] Non-PSO was also researched but was not factored into the NTG ratio because the energy savings could not be calculated accurately.

Table 5. Summary of Attribution Analysis

Utility	Sector	NTG	Free Ridership	Participant Spillover
Union	**Total**	**56%**	**54%**	**10%**
	Agriculture		0%	
	Commercial Retrofit		59%	
	Industrial		56%	
	Multifamily		42%	
	New Construction		33%	
Enbridge	**Total***	**79%**	**41%**	**21%**
	Agriculture		40%	
	Commercial Retrofit		12%	
	Industrial		50%	
	Multifamily		20%	
	New Construction		26%	

*Free ridership and spillover values include rounding error.
Source: Summit Blue Consulting. 2008. *Custom Projects Attribution Study.* Union Gas Limited and Enbridge Gas Distribution, October 27, 2008.

Following the study, the OEB approved the FR adjustment, but did not approve a spillover value. Currently, Union uses one NTG value for all C&I custom programs, the researched net-of-free ridership value calculated across all sectors (i.e., a FR of 54% and a net-of-free ridership value of 46%). Enbridge, on the other hand, currently uses the researched sector-specific net-of-free ridership values.

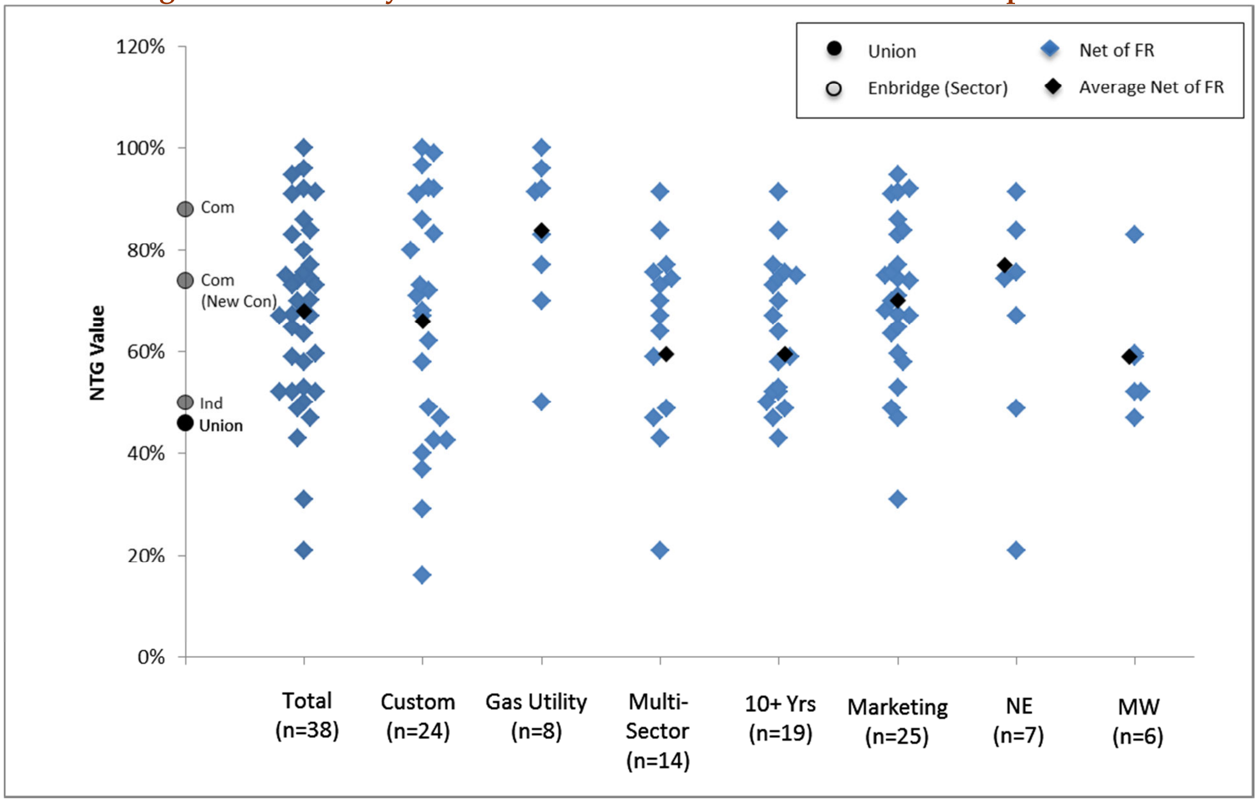
Comparing the current net-of-free ridership values for C&I custom programs (i.e., the researched net-of-free ridership values from the 2007-2008 Union and Enbridge study) to the range of researched values from the jurisdictional review provides context for the current net-of-free ridership values and insight into whether information available from other jurisdictions can be used to estimate NTG values in Ontario. Figure 14 summarizes findings from the review of researched NTG values in selected jurisdictions that are most relevant to Union and Enbridge.[21]

Union and Enbridge are gas-utilities that have been offering custom programs to commercial, industrial, or multi-sector customers for more than 10 years using both a direct marketing and channel/partner marketing strategy. As a result, Figure 14 presents the researched net-of-free ridership values for the following categories: custom program, gas utility, multi-sector, 10+

[21] We only summarize net-of-free ridership values as this summary provides the most information due to the largest sample sizes. Summaries of net of FR and spillover values are presented in Appendix E. Trends resulting from the jurisdictional review of NTG values that consider spillover should be interpreted with caution due to the small sample sizes.

years since program inception, a combination of direct and channel/partner marketing strategy, and northern regions (Northeast and Midwest).[22]

Figure 14. Summary of Relevant Researched Net-of-Free Ridership Values



Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG values (program-utility-year combinations).

The main findings resulting from the review of researched NTG values include the following:

- The NTG values calculated for Union and Enbridge are within the range of NTG values summarized in the review.
- When considering non-residential natural gas programs, NTG values appear to "cluster" between 40% and 90%. Union's NTG value is below the average. Enbridge's NTG value for the commercial sector is above the average while the NTG value for the industrial sector is below the average.

This "clustering" of values becomes less defined when considering other features of program design or implementation that make the NTG values more comparable to Union and Enbridge. For example, the clustering of NTG values for non-residential *custom* gas programs exhibits a wider dispersion without distinct clustering patterns.[23]

[22] All programs evaluated in the Midwest were offered in Wisconsin.
[23] Recall that when a NTG value was disaggregated into custom and prescriptive categories, these NTG values were included separately, resulting in more data points.

6. Assessing Options for NTG

Gross savings can usually be estimated quite accurately, however, estimating net savings poses greater challenges. Given the uncertainty around any NTG value, in this section Navigant applies a Decision Analysis approach for organizing information around alternative approaches to setting NTG values.

Navigant took the following steps to conduct the Decision Analysis:

1. Define the benefits of accurate (and costs of inaccurate) NTG values in a general context.
2. Narrow the focus the analysis on one of the benefits/cost for which Navigant had access to data; specifically, the incentives paid to utilities based on the estimated net savings (m^3) achieved.
3. Establish a baseline against which a sensitivity analysis can be conducted where a selected NTG value is assumed to be correct, but in fact is incorrect by some margin of error.[24]
4. Conduct a "value of information" analysis by examining the change in incentive payments resulting from better information on NTG values compared to the cost of obtaining the information (e.g., through NTG research).

This section concludes by organizing the results of the Decision Analysis to provide insight into the tradeoffs from using different approaches to setting an NTG value.

6.1 Decision Analysis

The first step in conducting the Decision Analysis is to identify the benefits resulting from more precise NTG values. Three of the primary benefits are described.

- **Program Design and Implementation.** NTG research can be leveraged to improve program design and implementation, ultimately providing greater gross and net savings. For example, FR research can inform decisions to discontinue incenting certain measures and boost the incentives for others. More generally, NTG research will identify what influences the customers' decisions regarding investments in energy efficiency, existing customer knowledge of energy efficiency and equipment operations, and identify aspects of the program that have the greatest influence on the customer's decision to participate in the program. NTG research can also provide insights into how the program is motivating distributors, contractors and other trade allies, and how their

[24] These first three steps are part of a "loss function" analysis which identifies the costs of selecting one NTG value when another value is the actual value. While a traditional loss function analysis focuses on deviations in both the mean value and the precision of the value, for simplicity, this analysis focuses only on precision or range of the values. Navigant did not conduct a more complex analysis because this simple approach provided insight into the value of more precise NTG values, i.e., a reduction in the range of NTG values.

actions might be leading to program spillover. All of this information helps in the design of improved programs.

- **Utility Incentive Payments.** Utilities, and utility shareholders, receive incentive payments for achieving performance goals. NTG values influence the incentive payments that are paid, or not paid, to utilities. More precise estimates of NTG values mitigate the risk that utilities face of receiving incentive payments that are too small, as well as the risk that ratepayers face of making incentive payments that are too large.[25]

- **Energy Savings as a Resource.** Regardless of the NTG value, the gross savings that result from the program are unchanged. (1) From a resource planning perspective, the net effects of the energy efficiency program must be known (i.e., the impacts attributable to the program must not have occurred in the absence of the program). (2) An accurate NTG estimate is important for understanding the equity implications of a program. I.e., participants that receive payments for taking actions that they would have taken even if the program had not existed transfers wealth from ratepayers to the participant. There are policy actions that can be taken to reduce equity issues, such as expanding the program to ensure all ratepayers have access to the program. However, a first step to considering the equity implications of a program is to accurately estimate the level of FR and spillover.

In the Decision Analysis that follows, Navigant focuses on the one benefit/cost for which data was available and for which there is little debate about how to formulate the benefit/cost: utility incentive payments. Union and Enbridge conducted an analysis of the sensitivity of utility incentive payments to changes in the NTG value of custom C&I programs.[26] The sensitivity analysis data was provided by the utilities and was not verified by Navigant.

6.1.1 Union

This section presents an assessment of the value of improved information on NTG values for Union Gas. Table 6 summarizes the impact on utility incentive payments if the custom NTG value is 10 percentage points higher or lower than the current custom NTG value of 0.46 used by Union.[27]

[25] While this report highlights the impact of improved precision of NTG values on the incentive payments received by the utilities, one can easily interpret the impact on ratepayers as it is a zero-sum game (i.e., the gain in incentive payments by utilities is a cost to ratepayers and vice versa).
[26] All other data inputs in the incentive payment calculations were held constant.
[27] This analysis assumes Union meets the targeted level of net savings.

NAVIGANT

Table 6. Value of Information Assessment for Union

	NTG Value for Custom Programs		Incentives	Change in Incentives
Base Case:	**Current NTG** NTG = 0.46	→	Incentives = $2.73 M	
Scenario 1:	**Higher True NTG** NTG = 0.56	→	Incentives = $5.63 M	**(+$2.90 M)**
Scenario 2:	**Lower True NTG** NTG = 0.36	→	Incentives = $0.8 M	**(-$1.93 M)**

Source: Sensitivity Analysis provided by Union.

At the net savings target under current assumptions, if the true custom program NTG value is 10 percentage points higher (Scenario 1) Union should receive an additional $2.9 million in incentive payments for savings achieved. If, instead, the true NTG value is 10 percentage points lower (Scenario 2), Union is receiving $1.93 million in incentives for savings that are not achieved.

A swing of +/- 10 percentage points (i.e., error bounds of +/- 22%) in the custom NTG value causes a swing in incentive payments by almost $3 million on the high side and $2 million on the low side. Assuming a revised custom program NTG value (e.g., by conducting NTG research) would reduce this margin of error by one-half, the error bounds would reduce to +/- 5 percentage points (i.e., +/- 11%) in the NTG value. The swing in incentive payments at the new error bounds would be approximately $1.5 million on the high side and $1 million on the low side. If the cost of revising the NTG values are less than $1 million given these assumed error bounds; then, revising the NTG values *could be judged to be warranted*.

6.1.2 Enbridge

This section presents an assessment of the value of improved information on NTG values for Enbridge. Table 7 summarizes the impact on utility incentive payments if the custom program NTG values are 10 percentage points higher or lower than the current custom NTG values used by Enbridge.[28]

[28] This analysis assumes Enbridge meets the targeted level of net savings.

Table 7. Value of Information Assessment for Enbridge

	NTG Value for Custom Programs		Incentives	Change in Incentives
Base Case:	**Current NTG by Program** Commercial = 0.80 Commercial New Construction = 0.74 Industrial = 0.50	→	Incentives = $2.58 M	
Scenario 1:	**Higher True NTG** Commercial = 0.90 Commercial New Construction = 0.84 Industrial = 0.60	→	Incentives = $4.26 M	**(+$1.68 M)**
Scenario 2:	**Lower True NTG** Commercial = 0.70 Commercial New Construction = 0.64 Industrial = 0.40	→	Incentives = $1.45 M	**(-$1.13 M)**

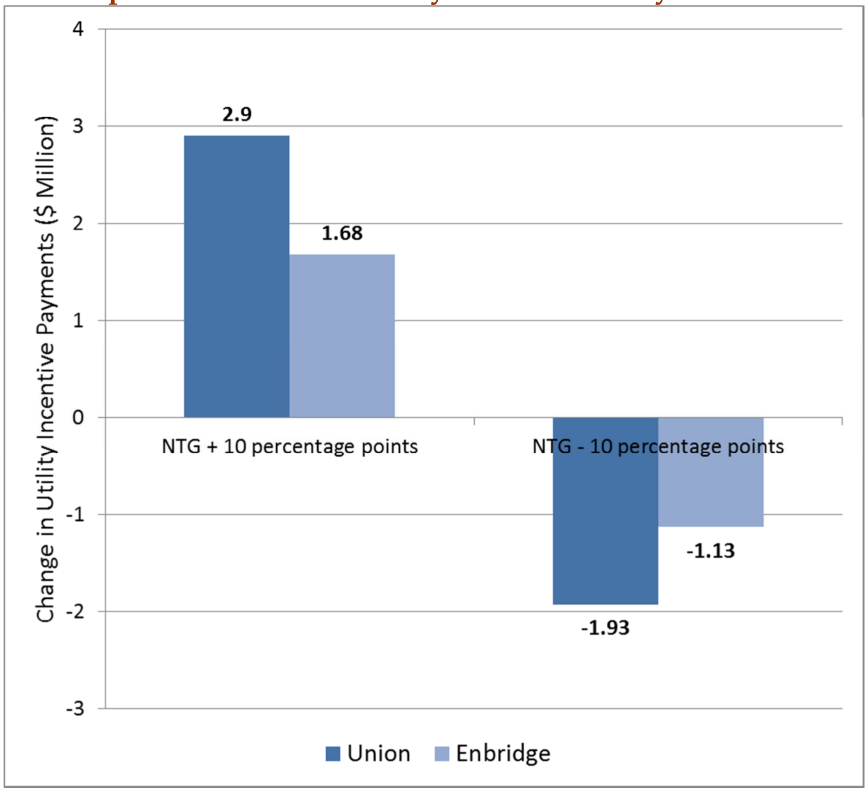
Source: Sensitivity Analysis provided by Enbridge.

At the net savings target under current assumptions, if the true custom program NTG values are 10 percentage points higher (Scenario 1) Enbridge should receive an additional $1.68 million in incentive payments for savings achieved. If, instead, the true custom program NTG values are 10 percentage points lower (Scenario 2), Enbridge is receiving $1.13 million in incentives for savings that are not achieved.

A swing of +/- 10 percentage points in custom program NTG values (i.e., error bounds of +/- 12.5% for commercial, +/- 13.5% for commercial new construction, and +/- 20% for industrial)) causes a swing in incentive payments by almost $2 million on the high side and $1 million on the low side. Assuming revised NTG values (e.g., by conducting NTG research) would reduce this uncertainty by one-half, the error bounds on the NTG values would reduce to +/- 5 percentage points in the NTG values. The swing in incentive payments at the new error bounds would be approximately $1 million on the high side and $0.5 million on the low side. If the cost of revising the NTG values are less than $0.5 million given these assumed error bounds; then, revising the NTG values *could be judged to be warranted*.

Figure 15 illustrates that the sensitivity in incentive payments to changes in custom program NTG values is greater for Union relative to Enbridge. This can be attributed to the fact that custom programs represent a larger share of Union's portfolio of programs, and consequently incentive payments, relative to Enbridge. Nevertheless, for both utilities changes in NTG values have a considerable impact on incentive payments.

Figure 15. Comparison of the Sensitivity of Incentive Payments to NTG Values



Source: Sensitivity analyses provided by Union and Enbridge.

6.2 *General Approaches to NTG*

In this section Navigant describes five general approaches to NTG representing the range of options for addressing net savings, from deeming a NTG value to conducting research to estimate a NTG value. The estimated increased precision of NTG values for each approach is identified as well as the approximate cost of the approach.

Option 1. Transfer NTG Values from Other Research

This approach transfers NTG values from the jurisdictional review. While the jurisdictional review revealed a wide range of NTG values, there is some clustering of values which could be used to inform a deemed value. If this approach is selected, the TEC sub-committee could select a NTG value from this clustering and apply it uniformly to Union and Enbridge's non-residential custom gas programs.

Advantages: The advantage of this approach is that it is simple, straightforward, uniform, and inexpensive.

Disadvantages: The disadvantage of this approach is that it does not recognize differences in the performance of different programs, designs, implementation, or program environments (such as economic conditions, energy prices, technology, and attitudes about climate change); consequently, the transferred values may provide inaccurate estimates of net savings.

<u>Option 2.</u> Adjusted or Scaled NTG Values based on Program Factors

This approach uses a simple scaled or adjusted NTG value from the jurisdictional review to better represent Union and Enbridge programs. A principal objective of the detailed review of researched NTG values was to summarize NTG values based on program factors comparable to Union and Enbridge programs. In particular, Navigant characterized researched NTG values by utility-type, program-type, targeted sector, program maturity, program marketing, and region. If this approach is selected, the TEC sub-committee could select a NTG value accounting for comparable program factors and adjusting appropriately for Union and Enbridge's non-residential custom gas programs. For example, a NTG value that includes spillover should be adjusted to reflect the fact that the majority of studies that consider spillover were conducted in recent years.

> *Advantages:* The advantage of this approach is that it is straightforward, uniform, and inexpensive. In addition, it recognizes differences in the performance of different program factors. Despite the disadvantages outlined below, the additional cost of adjusting or scaling the NTG value is so low that Option 2 is preferred in a pairwise comparison with Option 1.
>
> *Disadvantages:* The disadvantage of this approach is that due to the small number of researched NTG values with comparable program factors, the credibility of the scaled or adjusted NTG values may come into question, particularly if considering spillover.

<u>Option 3.</u> Align NTG Values using Limited Primary Data Collection

This approach augments comparative NTG values with a small set of selected primary data gathered during the course of program implementation and/or evaluation to enhance the precision of the NTG values. The detailed review revealed that in situations where program design remains consistent, NTG values can vary substantively from one program year to the next, likely due to changes in program implementation or program environment. Interviews with participating and non-participating trade allies, for example, can provide insight into FR and spillover, informing NTG values and requiring relatively limited data collection. If this approach is selected, the TEC sub-committee could select a comparable NTG value using limited primary data collection to adjust NTG values for Union and Enbridge's programs.

> *Advantages:* The advantage of this approach is that it recognizes differences in the performance of different programs, designs, implementation, and program environments while leveraging findings from the detailed review. NTG values will more accurately reflect actual net savings of the program.
>
> *Disadvantages:* One disadvantage may be the difficulty of developing the appropriate data to collect that represents actual changes in the NTG values. Another disadvantage

of this approach is that data collection, even if limited, can be costly; however, if it is incorporated within a program process, e.g., a short survey with the payment of incentives, the costs may be limited.

Option 4. Full NTG Research Study (After Program Year)

This approach conducts full-scale evaluations specific to Union and Enbridge programs at the end of the program-year cycle. There various methods for estimating net savings, including, for example, survey-based methods and econometric modeling. The enhanced self-report approach would likely be the most appropriate approach given Union and Enbridge's programs are custom C&I and that identifying the magnitude of individual NTG components is desired.

Advantages: The advantage of this approach is that it recognizes differences in the performance of different programs, designs, implementation, and program environments. Given a full-scale evaluation, NTG values will more accurately reflect actual net savings of the program relative to the limited data collection approach.

Disadvantages: The disadvantage of this approach is that full-scale evaluations are costly. In addition, if not designed properly, NTG research estimates may be biased. Appropriate NTG research contends with a variety of potential biases including, for example, non-response bias, recall bias, reaching the appropriate person, as well as biases related to respondents providing socially desirable responses or legitimizing past behavior.

Option 5. Integrated/Fast Feedback NTG Estimation

This approach relies on Integrated Data Collection, or rolling data collection processes, to estimate NTG values specific to Union and Enbridge programs using fast-feedback. Fast-feedback approaches reduce bias associated with NTG estimates, such as recall bias, by surveying participants closer to when the decision-making actually occurs (Energy Trust of Oregon 2012). Collecting data frequently over time assures that less biased estimates of FR are calculated.

Advantages: The advantage of this approach is that it recognizes differences in the performance of different programs, designs, implementation, and program environments. Integrated or Fast Feedback NTG estimation has received a lot of attention due to its ability to help address several key estimation issues – it is easier to target the appropriate people and recall bias is reduced by reducing the time cycle between project completion and data collection.[29] Another possible advantage of this approach is that program implementation staff can see what the NTG is as the program

[29] A number of recent studies estimating NTG make sure that they at least reach appropriate participating customers within 90 days after participating, and conduct surveys on a quarterly cycle. E.g., Summit Blue Consulting, LLC., Skumatz Economic Research Associates, Inc., and Quantec, LLC. 2005. *Commercial/Industrial Performance Program (CIPP) – Market Characterization, Market Assessment and Causality Evaluation.* NYSERDA, March 2005.

is implemented through the year. As a result, there are unlikely to be surprises in the NTG value at the end of a program year. Finally, this approach can actually be less costly than the traditional full research study presented above as Option 4 if data collection leverages existing program implementation efforts. For example, NTG surveys could be linked to the incentive payment process, e.g., one to two weeks after the incentives are paid a short free rider survey could be conducted (usually by phone). This approach is similar to Option 3 with more extensive data collection.

Disadvantages: The primary disadvantage of this approach are issues that may make integration difficult, e.g., appropriate timing of data collection, appropriate survey instruments, appropriate personnel leading the data collection all done along a timeline that is based on the implementation process. In addition, conducting research closer to program participation limits the amount of spillover that can be attributed to the program.

Table 8 provides a summary of the ability of the various approaches to improve the precision of the NTG value and provides an approximate cost of each NTG approach. Though an approximation, Navigant believe a 50% improvement in the precision of custom NTG values at a cost of $0.25 – 0.5 million is a reasonable estimate.[30]

Table 8. Ability of NTG Approaches to Produce More Precise NTG Values

General NTG Approach	Estimated Improved Precision (or Reduced Range) of NTG Value	Cost of NTG Approach per Utility (approximate)
Transfer NTG Values from Other Research	Little change	$3 – 5k
Adjust NTG Values based on Program Factors	Little change	$5 – 10k
Align NTG Values using Limited Primary Data	3 percentage points	$100 – 200k
Full NTG Research Study – After Program Year	5 percentage points	$250 – 500k
Integrated/Fast Feedback NTG Estimation	5 percentage points	$250 – 500k

Source: Navigant analysis.

[30] The cost estimates only reflect the contractor's program evaluation costs and do not include costs incurred by the utility and the TEC. These estimates assume primary data collection on program participants, a set of trade allies, and a sample of non-participants. Actual costs may vary depending on sub-strata and/or sector differentiation (e.g., commercial, commercial new construction, industrial).

7. Summary

The net savings of Union and Enbridge's custom C&I programs were first evaluated by Navigant (formerly Summit Blue Consulting) in 2007-2008. Following the study, the OEB approved the FR adjustment, but did not approve a spillover value. Since that time, there have been a host of program environment changes, including economic conditions, energy prices, advances in technology, as well as changes in the design and delivery of the custom programs. As a result, a key priority for Ontario's TEC sub-committee is to update the FR adjustment factor and reconsider the spillover adjustment.

As an initial step, the TEC sub-committee contracted Navigant to provide information to assist the TEC sub-committee in their deliberations on the appropriate approach to NTG for natural gas DSM programs in Ontario. Through a jurisdictional review of the approach to net savings, and a review of researched NTG values for programs comparable to Union and Enbridge's custom C&I gas programs, Navigant provides an assessment of the various approaches to NTG. Following is a summary of key findings:

Approach to Net Savings

- The majority of jurisdictions with ratepayer funded energy efficiency programs conduct NTG research, though only half adjust gross savings based on research.
- U.S. states that provide a performance incentive mechanism for utilities or program administrators are more likely to make deemed or researched NTG adjustments.
- There appears to be a trend towards considering participant and NPSO in NTG research in recent years.

Researched NTG Values in Selected Jurisdictions

- Navigant identified a total of 19 documents that conducted NTG research of non-residential gas programs that calculated 38 distinct results.
- Researched net-of-free ridership values for non-residential gas programs exhibit a wide dispersion (21% to 100%) with a slight "clustering" of values between 40% and 90%.
- Trends in researched NTG values that consider spillover, as well as trends when considering specific program characteristics, should be interpreted with caution due to the small sample sizes.
- Union and Enbridge's current NTG values are within the range of researched values. Union's NTG value is below the average value. Enbridge's NTG value for the commercial sector is above the average value while the NTG value for the industrial sector is below the average value.

- There are a variety of benefits of accurate (costs of inaccurate) NTG values that could be considered; utility incentive payments are just one.
- Improving the precision of NTG values has a sizable impact on incentive payments.
- NTG values with a margin of error of +/- 10 percentage points have roughly a $1 - $3 million impact on utility incentive payments.
- Alternate NTG approaches could improve the precision of NTG values by approximately 50% at an approximate cost of $0.25 - $0.50 million per utility.

The objective of this report is to provide information to assist the TEC sub-committee in their determination on the appropriate approach to NTG for DSM programs in Ontario, and not to provide a specific recommendation. While this report is not comprehensive in addressing all potential considerations, such as other benefits of accurate (costs of inaccurate) NTG values, it provides important information relevant to the discussion. In addition to summarizing the regulatory and methodological approach taken by other jurisdictions, and summarizing NTG values for programs with characteristics similar to Union and Enbridge's custom C&I programs, Navigant provides insight into the risks associated with inaccurate NTG values and the approximate cost of mitigating those risks.

Appendix A. General and Methodological References

Kushler, Martin, Nowak, Seth, and Patti White. 2012. *A National Survey of State Policies and Practices for the Evaluation of Ratepayer-Funded Energy Efficiency Programs.* American Council for an Energy-Efficient Economy. Available from: http://www.aceee.org/sites/default/files/publications/researchreports/u122.pdf.

MEEA. *Energy Efficiency Policies and Practices in Midwestern States.* Accessed January 23, 2013: http://mwalliance.org/policy/energy-efficiency-policies-and-practices-midwestern-states.

Messenger, Mike et al. 2010. *Review of Evaluation, Measurement and Verification Approaches Used to Estimate the Load Impacts and Effectiveness of Energy Efficiency Programs.* Lawrence Berkeley National Lab, April 2010. Available from: http://eetd.lbl.gov/ea/emp/reports/lbnl-3277e.pdf

Saxonis, W. 2007. *Free Ridership and Spillover: A Regulatory Dilemma.* Energy Program Evaluation Conference, Chicago, IL.

The Cadmus Group. 2012. *Assessment of Energy and Capacity Savings Potential in Iowa.* Iowa Utility Association, February 28, 2012.

The Cadmus Group. 2011. *Net-to-Gross: Updating Research.* Salt River Project, December 20, 2011.

Appendix B. Summary of NTG Values for Excluded Programs

There are two jurisdictions/programs that were excluded from the detailed review but provide additional information to the TEC sub-committee on NTG values in other jurisdictions.

California's **Savings by Design** program is a custom C&I program that has been offered for more than 10 years. This program was excluded from our review because the methodology used to calculate net savings was different from the approach used by the remaining documents reviewed. In particular, responses to a FR survey were used to adjust the baseline of an engineering model. The NTG ratio was then calculated as the ratio of gross to net savings, as estimated by the engineering model. This approach accounts for interactive effects between measures and resulted in NTG values greater than 100%, even though only a FR adjustment was made. The table below summarizes the NTG values for Savings by Design.

NTG Values for Savings by Design

Category	NTG Value
Combined	87%
PG&E	66%
SDG&E	109%
SCE	101%
SCG	25%

Source: RLW Analytics. 2008. *An Evaluation of the 2004-2005 Savings by Design Program.* California Public Utilities Commission, October 2008.

NYSERDA has implemented a number of C&I programs with custom components, and include both electric and gas measures. Relevant programs include: **Industrial and Process Efficiency, Flexible Technical Assistance, C&I Performance, and New Construction Program.** Recent research estimates NTG values using a rigorous methodology, but were excluded from our review because the values were not reported separately for electric and gas measures. The Table below summarizes NTG values for these programs, where NTG = 1 – Free Ridership + Participant Spillover + Non-Participant Spillover.

NTG Values for NYSERDA Programs

Program	NTG Value
Industrial and Process Efficiency	104%
Flexible Technical Assistance	117%
New Construction Program	116%
C&I Performance	123%

Sources: Megdal & Associates. 2012. *NYSERDA 2009-2010 Industrial and Process Efficiency Program Impact Evaluation Report; Impact Evaluation: NYSERDA 2007-2009 FlexTech Program; New Construction Program (NCP) Impact Evaluation Report for Program Years 2007-2008;*
Summit Blue Consulting. 2007. *Commercial and Industrial Performance Program (CIPP): Market Characterization, Market Assessment and Causality Evaluation.* NYSERDA, May 2007.

NAVIGANT

Appendix C. Annotated Bibliography of Documents Reviewed

2004/2005 Statewide Express Efficiency and Upstream HVAC Program Impact Evaluation	
Author and Date	Itron and KEMA. December 31, 2008.
Jurisdiction	California
Utilities	Pacific Gas & Electric, San Diego Gas & Electric, Southern California Edison, and Southern California Gas Company
Program Name	Express Efficiency Program
Program Summary	The Express Efficiency program targets small and medium-sized commercial customers (electricity demand less than 500 kW; annual gas consumption less than 250,000 therms) providing financial incentives to end-users for the installation of selected energy efficient electric and gas technologies (e.g., lighting, refrigeration, air conditioning, food service, agricultural, and gas technologies). The program implements a marketing strategy directly with the end-user and through upstream partners (e.g., vendors).
Program Year	2004-2005
NTG	0.51
Free-Ridership	NTG=1-FR; 0.49
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Self-report. Participant surveys were completed by end-users. The free-ridership score was the average of scores from two methodologies using participant survey data. One methodology adjusts for timing.

Note that this evaluation study also addresses the Upstream HVAC/Motors; however, no gas savings were reported under this program in 2004-2005.

2004-2005 Statewide Nonresidential Standard Performance Contract Program Measurement and Evaluation Study	
Author and Date	Itron. September 30, 2008.
Jurisdiction	California
Utilities	Pacific Gas & Electric, San Diego Gas & Electric, and Southern California Edison
Program Name	Nonresidential Standard Performance Contract Program
Program Summary	This program provides financial incentives for custom cost effective energy saving retrofits of existing facilities. While targeted at large and medium-sized businesses, small businesses can participate if they are ineligible for incentives through California's Express Efficiency program. Major measure types include: lighting and lighting controls, variable speed-drive for motors, HVAC, and industrial processes. Pacific Gas & Electric and San Diego Gas & Electric offer incentives for energy efficiency gas measures, with incentives of $1.00 per therm.
Program Year	2004-2005
NTG	0.57
Free-Ridership	0.43
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Self-report. Participant surveys were completed by end-users. The sample used for gross impact analysis was also used for net impact analysis. The free-ridership score was the average of scores from two methodologies using participant survey data, in which one methodology adjusted for timing.

2006-2008 Retro-Commissioning Impact Evaluation	
Author and Date	SBW Consulting. February 8, 2010.
Jurisdiction	California
Utilities	Pacific Gas & Electric, San Diego Gas & Electric, Southern California Edison, and Southern California Gas
Program Name	More than two dozen Retro-Commissioning programs.
Program Summary	This report presents evaluation, measurement and verification activities for over two dozen commercial retro-commissioning programs that target high impact measures (i.e. contribute more than 1% of utilities' savings portfolio). Given the number of programs, program design varies and may include technical assistance and/or financial incentives.
Program Year	2006-2008
NTG	PG&E: 0.86 SCE: 0.91 SCG: 0.92 SDG&E: 0.68
Free-Ridership	PG&E: 0.14 SCE: 0.09 SCG: 0.08 SDG&E: 0.32
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Enhanced self-report. Includes participant surveys, vendor surveys, program staff interviews, and file reviews. In some cases supplemental questions were asked of participant decision-makers. Free-ridership estimate is based on survey questions about timing and selection, program influence, and likelihood. Timing adjustments are included. When multiple elements feed into one score, the maximum (representing highest program influence) is used.

2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group	
Author and Date	Itron. February 3, 2010.
Jurisdiction	California
Utilities	Pacific Gas & Electric
Program Name	Program administered by PG&E: • Fabrication, Process and Manufacturing Programs administered by a third-party: • Heavy Industry Energy Efficiency Program • California Wastewater Process Optimization Program • Energy Efficiency Services for Oil Production • Wastewater Process Efficiency Initiative • Refinery Energy Efficiency Program • Assessment, Implementation and Monitoring • Value and Energy Stream Mapping Advantage Plus • Energy Efficiency of Compressed Systems • C&I Boiler Efficiency Program
Program Summary	The Pacific Gas & Electric Fabrication, Process and Manufacturing contract group is comprised of one PG&E program and nine third-party programs. These programs provide technical assistance and financial incentives for the installation of custom and prescriptive electric and gas measures in industrial facilities. Eligible sectors include industrial and manufacturing, water supply and treatment, wastewater, oil and gas extraction, refining, and production. Major measure types include: boiler upgrades and controls, boiler heat recovery, pipe and duct insulation, HVAC, process improvements, as well as various electric measures.
Program Year	2006-2008
NTG	0.31
Free-Ridership	0.69
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Enhanced self-report. Includes participant surveys, vendor surveys, program staff interviews, and file reviews. In some cases supplemental questions were asked of participant decision-makers. Free-ridership estimate is based on survey questions about timing and selection, program influence, and likelihood. Timing adjustments are included. When multiple elements feed into one score, the maximum (representing highest program influence) is used.

NAVIGANT

Evaluation of the Southern California Gas Company 2004-2005 Non-Residential Financial Incentives Program	
Author and Date	ECONorthwest. June 6, 2006.
Jurisdiction	California
Utilities	Southern California Gas Company
Program Name	Nonresidential Financial Incentives Program
Program Summary	This program provides technical assistance, education, and financial incentives for prescriptive and custom energy efficiency gas measures. This program is targeted at small and medium-sized customers, spanning the commercial, industrial and agricultural sectors. There are three program offerings: • The Commercial Food Service Equipment Rebate program offering provides financial incentives for prescriptive measures. Examples include ovens, broilers, griddles, and fryers. • The Nonresidential Equipment Replacement program offering provides financial incentives for the replacement of existing gas technologies with energy efficient alternative. Examples include industrial furnaces, ovens, dryers, washers, and more. • The Nonresidential Energy Conservation program offering provides financial incentives for energy efficiency retrofits and energy efficiency improvements to industrial processes. Examples include heat-recovery, process steam improvements, and high-efficiency burner replacements.
Program Year	2004-2005
NTG	0.70
Free-Ridership	0.30
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Self-report. Participant surveys were completed by end-users. Three methodologies were implemented though a preferred methodology is identified. This methodology calculates a probability of influence based on the influence of the financial incentive, program representatives, and adjusts for timing.

Comprehensive Process and Impact Evaluation of the Business Heating Efficiency Program - Colorado	
Author and Date	TetraTech. December 14, 2011.
Jurisdiction	Colorado
Utilities	Xcel Energy
Program Name	Business Heating Efficiency Program
Program Summary	This program provides financial incentives to commercial customers for prescriptive energy efficient gas measures. Major measure types include: new high efficiency hot water boilers and furnaces, improvements to existing boilers and hot water heaters, or boiler tune-ups to maintain peak operating efficiency.
Program Year	2011
NTG	0.85
Free-Ridership	0.26
Participant Spillover	0.11 (Like)
Non-Participant Spillover	N/A – Conducted interviews with HVAC trade allies but were unable to quantify NPSO.
Research Method	Self-report. Surveys include questions about the timing and selection of program measures, the influence of the program (whether rebate, recommendation, or other program intervention), and the likelihood of various actions now and in the future had the program not been available. Methodology adjusts free-ridership score if past program participation in any Xcel Energy program influences the decision to install a measure. Spillover is considered if it occurs within 4 years.

NAVIGANT

2011 C&I Natural Gas Programs Free-Ridership and Spillover Study	
Author and Date	TetraTech. June 26, 2012.
Jurisdiction	Massachusetts
Utilities	National Grid, NSTAR, Unitil, Berkshire Gas, Columbia Gas, and New England Gas
Program Names	All C&I custom and prescriptive gas programs were included in this evaluation. • National Grid programs include: New Construction (custom and prescriptive), Retrofit (custom and prescriptive), Direct Install (prescriptive) • NSTAR programs include: Business Solutions (custom), Construction Solutions (custom), Small Business Solutions (custom and prescriptive) • Columbia Gas programs include: Large Custom, Small Custom, Prescriptive • Unitil programs include: Large Retrofit (custom and prescriptive), Gas Networks (prescriptive), Small Direct Install (prescriptive) • New England Gas programs include: Retrofit (custom), Lost Opportunity (prescriptive), Direct Install (prescriptive) • Berkshire Gas programs include: Custom, Prescriptive
Program Summary	These programs provide financial incentives for installing custom and prescriptive energy efficient gas measures.
Program Year	2011
NTG	0.79
Free-Ridership	0.305
Participant Spillover	0.085 (Like)
Non-Participant Spillover	0.007
Research Method	Enhanced self-report. Combination of participant (decision-makers) and trade ally surveys. Surveys include questions about likelihood of equivalent efficiency and quantity of program measures, as well as the timing. Questions were also included about the influence of program and various features of the program, as well as the influence of participating in past programs. Free-ridership and spillover estimates are weighted by therm savings and the probability of being surveyed. Surveys with design professionals and equipment vendors were used to calculate free-ridership in cases where the decision was heavily influenced by the design professional/equipment vendor, as well as to calculate NPSO.

NAVIGANT

Achieving Natural Gas Savings Goals: Commercial Heating Programs Heat It Up	
Author and Date	TetraTech and Xcel Energy. 2012 ACEEE Summer Study on Energy Efficiency in Buildings.
Jurisdiction	Minnesota
Utilities	Xcel Energy
Program Name	Business Heating Efficiency Program
Program Summary	This program provides financial incentives to commercial customers for prescriptive energy efficient gas measures. Major measure types include: new high efficiency hot water boilers and furnaces, improvements to existing boilers and hot water heaters, or boiler tune-ups to maintain peak operating efficiency.
Program Year	2011
NTG	1.09
Free-Ridership	0.17
Participant Spillover	0.26 (Like)
Non-Participant Spillover	N/A – Conducted interviews with HVAC trade allies but were unable to quantify NPSO.
Research Method	Self-report. Surveys include questions about the timing and selection of program measures, the influence of the program (whether rebate, recommendation, or other program intervention), and the likelihood of various actions now and in the future had the program not been available. Methodology adjusts free-ridership score if past program participation in any Xcel Energy program influences the decision to install a measure. Spillover is considered if it occurs within 4 years.

Note: Research method is the method employed by TetraTech in the evaluation of Colorado's Xcel Energy Business Heating Efficiency Program which is the same method employed in Minnesota. This paper relies on TetraTech's evaluation to report NTG values, though the report itself is not publicly available.

New Jersey's Clean Energy Program Energy Impact Evaluation: SmartStart Program Impact Evaluation	
Author and Date	KEMA. September 17, 2009.
Jurisdiction	New Jersey
Utilities	New Jersey's Clean Energy Program
Program Name	SmartStart Buildings Program (New Construction, Schools, and Retrofit program)
Program Summary	This program provides financial incentives and technical assistance for energy efficient measures in new construction, retrofits of existing buildings, and schools.
Program Year	2006
NTG	0.21
Free-Ridership	0.79
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Self-report. Surveys include questions about likelihood of equivalent efficiency and quantity of program measures, as well as the timing. Free-ridership measures for timing, efficiency, and quantity are multiplied to determine free-ridership. Adjustments to free-ridership score based on timing is made. The sample size for Schools and New Construction programs is small.

Evaluation of 2011 DSM Portfolio	
Author and Date	ADM Associates. June 29, 2012.
Jurisdiction	New Mexico
Utilities	New Mexico Gas Company
Program Names	Commercial Solutions, Commercial High Efficiency Water Heater, Commercial Energy Star Food Service, and SCORE Pilot
Program Summary	These programs provide financial incentives for custom and prescriptive measures installed by commercial customers. • The Commercial Solutions program includes two program offerings: direct install of low flow faucet aerators and pre-rinse spray valves, and custom incentives of up to $0.75 per therm for custom measures, such as: water heating, HVAC, building envelope, and industrial processes. The SCORE Pilot is similar to the Commercial Solutions program but is targeted at K-12 school districts. • The Commercial Energy Star Food Services program provides prescriptive rebates for commercial kitchen measures, such as fryers, dishwashers, convection ovens, and commercial griddles. • The Commercial High Efficiency Water Heater program provides financial incentives for storage tank and tankless water heaters.
Program Year	2011
NTG	Commercial Solutions: 0.96 Commercial High Efficiency Water Heater: 1.00 Commercial Energy Star Food Service: 1.00 SCORE Pilot: 1.00
Free-Ridership	Commercial Solutions: 0.04
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Self-report. Surveys include questions about the financial ability to purchase measures without the program, the importance of the financial incentive, prior planning to purchase measures, and demonstrated behavior in purchasing similar measures without a financial incentive.

Fast Feedback Results	
Author and Date	Energy Trust of Oregon. April 25, 2012.
Jurisdiction	Oregon
Utilities	Energy Trust of Oregon
Program Names	Existing Buildings Program, Production Efficiency Program
Program Summary	Descriptions of programs not included in study. Information that follows is from the Energy Trust of Oregon's website (http://energytrust.org) Existing Buildings program provides custom and prescriptive financial incentives to existing commercial facilities. Major gas measure types include: HVAC, furnace, radiant heater, hot water tanks, tankless water heaters, boilers, and steam traps. Production Efficiency program provides technical assistance and financial incentives for energy efficiency improvements for industrial processes, including manufacturing, agriculture, and water/wastewater treatment. Major measure types include: motors, compressed air, variable speed drives, refrigeration, pumps, fans, and lighting.
Program Year	Q2 2010
NTG	Existing Buildings: 0.73 Existing Multifamily: 0.52 Production Efficiency: 0.80
Free-Ridership	Existing Buildings: 0.27 Existing Multifamily: 0.48 Production Efficiency: 0.20
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Self-report. Surveys are conducted with participants that received a financial incentive within the previous month. The survey is designed to be completed in no more than 5 minutes and consists of 10 questions or less. Free-ridership is calculated as the sum of a project change score and an influence score. The project change score is based on survey questions about the actions the customer would have taken if the program was not available. Influence questions ask about the influence of the program, trade ally influence, etc.

Impact and Process Evaluation of the 2006-2007 Building Efficiency Program	
Author and Date	Research Into Action and the Cadmus Group. August 3, 2009.
Jurisdiction	Oregon
Utilities	Energy Trust of Oregon
Program Name	Building Efficiency Program
Program Summary	This program provides technical assistance and financial incentives for electric and gas energy-saving measures installed by commercial and institutional customers. Financial incentives are provided for both prescriptive and custom measures. Major measure types include: lighting, motors, HVAC, gas space and water heaters, restaurant equipment, and insulation.
Program Year	2006-2007
NTG	0.70
Free-Ridership	0.30
Participant Spillover	Qualitative assessment.
Non-Participant Spillover	N/A
Research Method	Self-report. Survey questions consider program influence, intentions for the project without the program, and budget.

Evaluation of Building Efficiency Program 2004 &2005	
Author and Date	ADM Associates. February 2009.
Jurisdiction	Oregon
Utilities	Energy Trust of Oregon
Program Name	Building Efficiency Program
Program Summary	This program provides technical assistance and financial incentives for electric and gas energy-saving measures installed in existing commercial, institutional, and agricultural facilities. Financial incentives are provided for both prescriptive and custom measures. Major measure types include: lighting, motors, HVAC, gas space and water heaters, restaurant equipment, and insulation.
Program Year	2004-2005
NTG	2004: 0.65 2005: 0.95
Free-Ridership	2004: 0.35 2005: 0.05
Participant Spillover	Qualitative assessment.
Non-Participant Spillover	N/A
Research Method	Self-report. Survey questions consider program influence, intentions for the project without the program/prior planning, and previous experience with the measure. Each question is binary (i.e. yes/no). Partial free-ridership is explored through questions about efficiency level, quantity and timing.

Impact and Process Evaluation of the 2006-2007 New Building Efficiency Program	
Author and Date	ADM Associates. June 2009.
Jurisdiction	Oregon
Utilities	Energy Trust of Oregon
Program Name	New Building Efficiency Program
Program Summary	This program provides technical assistance and financial incentives for electric and gas energy-saving measures installed in new commercial facilities or commercial facilities undergoing major renovation. Major measure types include: lighting, HVAC, motors, energy management systems, and washer/dryers.
Program Year	2006-2007
NTG	0.67
Free-Ridership	0.33
Participant Spillover	Qualitative assessment.
Non-Participant Spillover	N/A
Research Method	Self-report. Participant surveys were conducted. Free-ridership estimates are based on survey questions that ask about the influence of the program, the participants' intentions for the project if the program were not available, and their financial ability to install the measures if the program were not available.

C&I Energy Efficiency Retrofit Custom Programs Portfolio Evaluation	
Author and Date	Navigant Consulting. February 3, 2012.
Jurisdiction	Washington
Utilities	Puget Sound Energy
Program Name	Custom Grant Program
Program Summary	This program provides financial incentives for the installation of custom energy efficient measures as part of a retrofit, new construction, or expansion of existing facilities project. Major measure types include: lighting, boilers, HVAC, variable speed drives, and process improvements.
Program Year	2010-2011
NTG	1.02-1.1
Free-Ridership	0.27
Participant Spillover	0.07-0.09 (inside like); 0.04-0.05 (outside like)
Non-Participant Spillover	0.18-0.23
Research Method	Self-report. Surveys of participants and non-participants were conducted. Free-ridership was estimated based on survey questions about timing, efficiency, quantity, and program importance. Spillover calculated as a factor of savings derived from spillover project based on program influence. Savings were assumed equal to savings by in-program projects (by measure-type). Similar calculations were conducted for NPSO.

Focus on Energy Evaluation: Business Programs – Additional Looks at Attribution	
Author and Date	PA Consulting Group and KEMA. February 26, 2010.
Jurisdiction	Wisconsin
Utilities	Focus on Energy
Program Name	The names of specific program offerings are not reported.
Program Summary	Various programs provide technical assistance and financial incentives for implementing cost effective energy efficiency measures. Both prescriptive and custom incentives are available. Targeted sectors include commercial, industrial, agricultural, and institutional. Major measure types include: boilers, HVAC, refrigeration, water heater, expanded processes, and lighting.
Program Year	July 1, 2007 through September 30, 2008
NTG	0.52
Free-Ridership	0.48
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Enhanced self-report. Surveys of participants and trade allies were conducted. Free-ridership survey questions ask about timing, efficiency, and the quantity of measures installed if the program were not available. These free-ridership estimates are multiplied (e.g., NTG=1-FqFeFt). Surveys include consistency checks. NTG estimates based on participant survey data is compared to estimates based on trade ally survey data. The maximum value is selected.

NAVIGANT

Focus on Energy Evaluation: Business Programs Impact Evaluation Report – Last Quarter of Calendar Year 2009 and First Two Quarters of Calendar Year 2010	
Author and Date	TetraTech and KEMA. January 27, 2011.
Jurisdiction	Wisconsin
Utilities	Focus on Energy
Program Name	The names of specific program offerings are not reported.
Program Summary	Various programs provide technical assistance and financial incentives for implementing cost effective energy efficiency measures. Both prescriptive and custom incentives are available. Targeted sectors include commercial, industrial, agricultural, and institutional. Major measure types include: boilers, HVAC, refrigeration, water heater, expanded processes, and lighting.
Program Year	October 1, 2009 through June 30, 2010
NTG	2009: 0.60 2010: 0.47
Free-Ridership	2009: 0.40 2010: 0.53
Participant Spillover	(Identified in a separate study as 0.002%)
Non-Participant Spillover	N/A
Research Method	Enhanced self-report. Surveys of participants and trade allies were conducted. Free-ridership survey questions ask about timing, efficiency, and the quantity of measures installed if the program were not available. These free-ridership estimates are multiplied (e.g., NTG=1-FqFeFt). Surveys include consistency checks. NTG estimates based on participant survey data is compared to estimates based on trade ally survey data. The maximum value is selected.

Focus on Energy Evaluation: Semiannual Report (Second Half of 2009)	
Author and Date	PA Consulting Group. April 23, 2010.
Jurisdiction	Wisconsin
Utilities	Focus on Energy
Program Name	The names of specific program offerings are not reported.
Program Summary	Various programs provide technical assistance and financial incentives for implementing cost effective energy efficiency measures. Both prescriptive and custom incentives are available. Targeted sectors include commercial, industrial, agricultural, and institutional. Major measure types include: boilers, HVAC, refrigeration, water heater, expanded processes, and lighting.
Program Year	Q3 and Q4 2009
NTG	0.59
Free-Ridership	0.41
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Enhanced self-report. Participant surveys and surveys with trade allies were conducted. Free-ridership survey questions ask about timing, efficiency, and the quantity of measures installed if the program were not available. Conducted a sensitivity analysis on treatment of timing using methodologies adopted in other jurisdictions finding little variation.

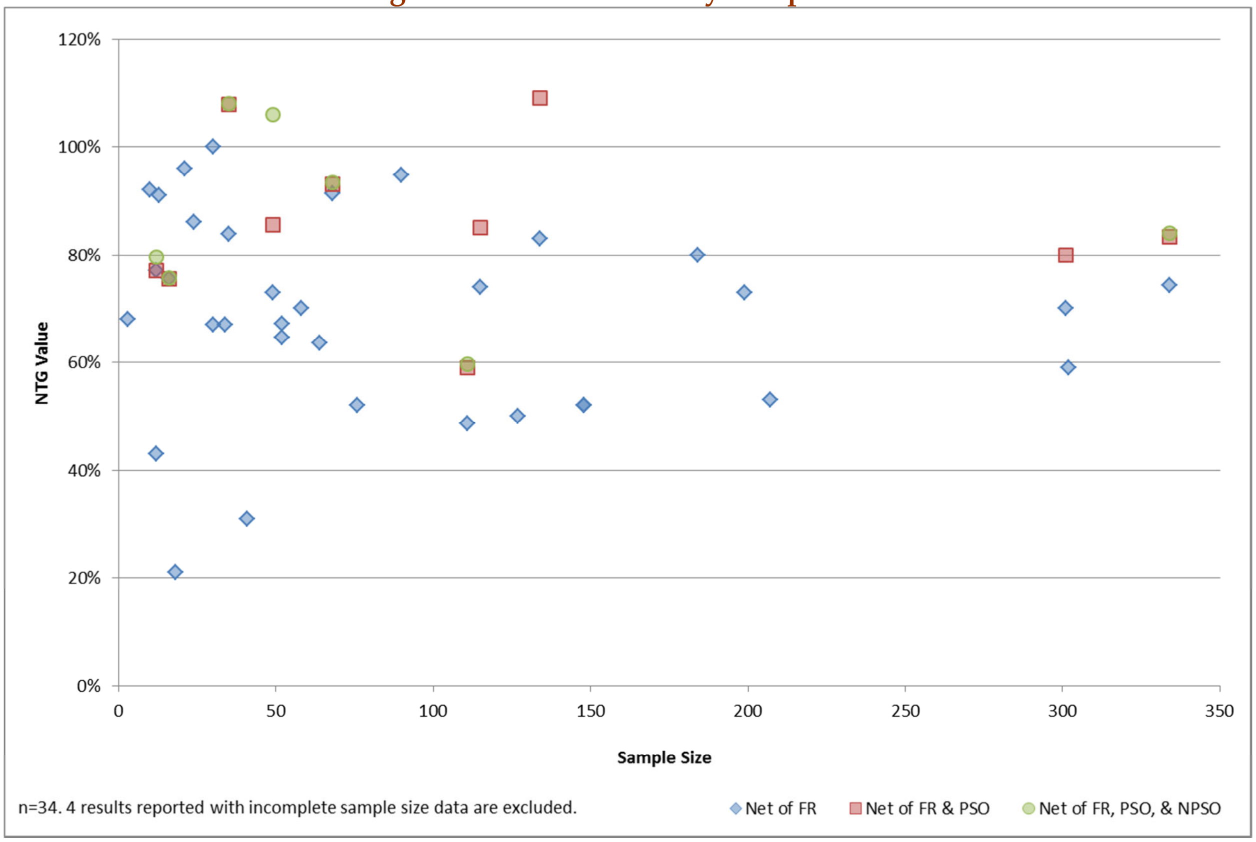
Focus on Energy Evaluation: Semiannual Report (First Half of 2009)	
Author and Date	PA Consulting Group. October 19, 2009.
Jurisdiction	Wisconsin
Utilities	Focus on Energy
Program Name	The names of specific program offerings are not reported.
Program Summary	Various programs provide technical assistance and financial incentives for implementing cost effective energy efficiency measures. Both prescriptive and custom incentives are available. Targeted sectors include commercial, industrial, agricultural, and institutional. Major measure types include: boilers, HVAC, refrigeration, water heater, expanded processes, and lighting.
Program Year	A1 and A2 2009
NTG	0.52
Free-Ridership	0.48
Participant Spillover	N/A
Non-Participant Spillover	N/A
Research Method	Enhanced self-report. Participant surveys and surveys with trade allies were conducted. Free-ridership survey questions ask about timing, efficiency, and the quantity of measures installed if the program were not available.

NAVIGANT

Appendix D. NTG Values by Sample Size

The figure below summarizes NTG values by sample size. Sample sizes are reported in raw form and do not reflect the percent of participants or percent of energy savings. Consequently, this Figure should be interpreted with caution.

Figure D1. NTG Values by Sample Size

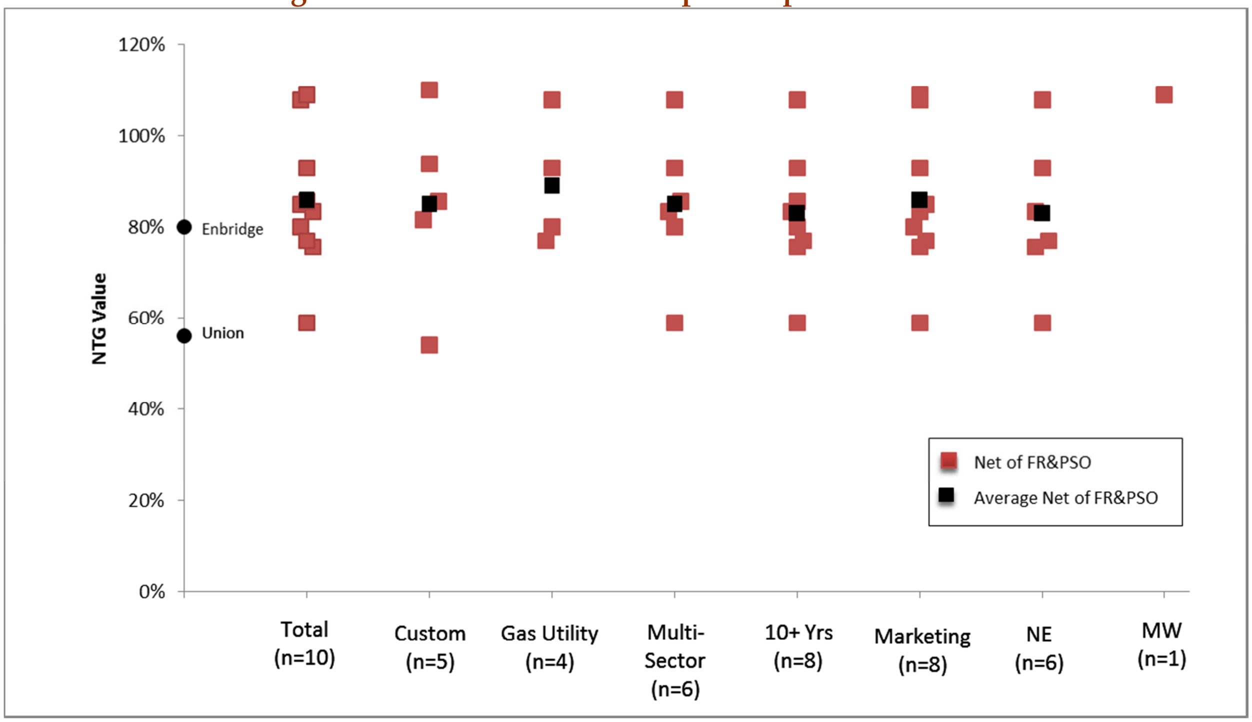


Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG results (program-utility-year combinations) reported in the 19 studies.

Appendix E. Researched Net-of-Free Ridership and Spillover Values

The figure below summarizes net-of-free ridership and PSO values that are most relevant to Union and Enbridge programs. In particular, values are presented for the following categories: custom program, gas utility, multi-sector, 10+ years since program inception, a combination of direct and channel/partner marketing strategy, and northern regions (Northeast and Midwest). Note that the values reported for Union and Enbridge are researched values representing all sectors resulting from the 2007-2008 attribution study. Caution should be used in interpreting trends due to the small sample sizes. Nevertheless similar trends emerge. Enbridge and Union NTG values are below the average values.

Figure E1. Net-of-Free Ridership and Spillover Values



Source: Navigant analysis. Note that the sample size (n) represents the number of unique NTG results (program-utility-year combinations).