Hydro One Networks Inc.

EB-2019-0082

OEB Staff Compendium

Panel 4

November 1, 2019

APPENDIX F

FORECAST ACCURACY

Tables 6a to 6c present the forecast accuracy of the OEB-approved forecasts of the three charge determinants on a weather-corrected basis for the past six rate applications (EB-2006-0501, EB-2008-0272, EB-2010-0002, EB-2012-0031, EB-2014-00140, and EB-2016-0160).

All forecasts are weather-normal and compared with weather-corrected actuals. In all tables, a negative or positive percent deviation indicates that the forecast was below or above actual-weather corrected.

Table 6a
Historical Board Approved for Network Connection Forecast
vs. Historical Actual and Historical Actual-Weather Normalized

	12-Month Average in MW								Difference from Actual Weather Corrected (%)					
Year	EB-2006-0501 Forecast (1)	EB-2008-0272 Forecast (2)	EB-2010-0002 Forecast (3)	EB-2012-0031 Forecast (4)	EB-2014-0140 Forecast (5)	EB-2016-0160 Forecast (6)	Actual: Weather Corrected	Actual	EB-2006-0501 Forecast	EB-2008-0272 Forecast	EB-2010-0002 Forecast	EB-2012 0031 Forecast	EB-2014-0140 Forecast	EB-2016-0160 Forecast
2005	21,704						21,702	22,507	0.01					
2006	21,259						21,275	22,028	-0.08					
2007	20,827	20,928					20,928	22,398	-0.48	0.00				
2008	20,872	20,943					21,067	21,307	-0.92	-0.59				
2009		20,842	20,868				20,868	20,410		-0.13	0.00			
2010		20,199	20,414				20,330	21,196		-0.64	0.41			
2011			20,150	20,245			20,245	20,861			-0.47	0.00		
2012			19,845	20,042			20,086	20,868			-1.20	-0.22		
2013				20,023	20,220		20,220	21,352				-0.97	0.00	
2014				19,552	20,276		20,601	20,643				-5.09	-1.58	
2015					20,559	20,236	20,236	20,384					1.60	0.00
2016					20,779	20,265	20,245	20,630					2.64	0.10
2017						20,405	19,705	19,608						3.55
2018						20,410	19,678	20,585						3.72
Average Excluding First Year (Actual) (7)									-0.49	-0.45	-0.42	-2.10	0.89	2.46

(1) Forecast: EB-2006-0501; Ex A; T14; S 3; P 19 of 20.
(2) Forecast: EB-2008-0272; Ex A; T14; S 3; P 22 of 24.
(3) Forecast: EB-2010-0002; Ex A; T14; S 3; P 19 of 21.
(4) Forecast: EB-2012-0031; Ex A; T15; S 2; P 22 of 24.
(5) Forecast: EB-2014-0140; Ex A; T15; S 2; P 20 of 23, settlement amount shown.
(6) Forecast: EB-2016-0160; Ex E1; T3; S 1; P 20 of 52.
(7) Compares actual-weather corrected with forecast (3 years of forecast for EB-2006-0501, EB-2008-0272, EB-2010-0002, EB-2012-0031, EB-2014-0140, and EB-2016-0160 forecast).

Table 6b
Historical Board Approved for Line Connection Forecast
vs. Historical Actual and Historical Actual-Weather Normalized

Year	12-Month Average in MW: EB-2006-0501 Forecast (1)	EB-2008-0272 Forecast (2)	EB-2010-0002 Forecast (3)	EB-2012-0031 Forecast (4)	EB-2014-0140 Forecast (5)	EB-2016-0160 Forecast (6)	Actual: Weather Corrected	Actual	Difference from Actual Weather Corrected (%) (5): EB-2006-0501 Forecast	EB-2008-0272 Forecast	EB-2010-0002 Forecast	EB-2012 0031 Forecast	EB-2014-0140 Forecast	EB-2016-0160 Forecast
2005	20,590						20,590	21,345	0.00					
2006	20,242						20,282	20,991	-0.20					
2007	19,875	20,044					20,044	21,443	-0.84	0.00				
2008	19,940	20,111					20,156	20,386	-1.07	-0.23				
2009		20,100	19,796				19,796	19,372		1.53	0.00			
2010		19,555	19,674				19,348	20,162		1.07	1.69			
2011			19,500	19,417			19,417	20,004			0.42	0.00		
2012			19,286	19,359			19,298	20,047			-0.06	0.32		
2013				19,406	19,322		19,322	20,405				0.44	0.00	
2014				18,990	19,488		19,626	19,843				-3.24	-0.70	
2015					19,851	19,576	19,576	19,829					1.40	0.00
2016					20,150	19,605	19,540	20,027					3.12	0.33
2017						19,741	19,100	19,064						3.35
2018						19,746	19,137	20,040						3.18
Average Excluding First Year (Actual) (7)									-0.71	0.79	0.68	-0.83	1.27	1.84

(1) Forecast: EB-2006-0501; Ex A; T14; S 3; P 19 of 20.
(2) Forecast: EB-2008-0272; Ex A; T14; S 3; P 22 of 24.
(3) Forecast: EB-2010-0002; Ex A; T14; S 3; P 19 of 21.
(4) Forecast: EB-2012-0031; Ex A; T15; S 2; P 22 of 24.
(5) Forecast: EB-2014-0140; Ex A; T15; S 2; P 20 of 23, settlement amount shown.
(6) Forecast: EB-2016-0160; Ex E1; T3; S 1; P 20 of 52.
(7) Compares actual-weather corrected with forecast (3 years of forecast for EB-2006-0501, EB-2008-0272, EB-2010-0002, EB-2012-0031, EB-2014-0140, and EB-2016-0160 forecast).

Table 6c
Historical Board Approved for Transforer Connection Forecast
vs. Historical Actual and Historical Actual-Weather Corrected

	12-Month Average in MW								Difference from Actual Weather Corrected (%) (5)					
Year	EB-2006-0501 Forecast (1)	EB-2008-0272 Forecast (2)	EB-2010-0002 Forecast (3)	EB-2012-0031 Forecast (4)	EB-2014-0140 Forecast (5)	EB-2016-0160 Forecast (6)	Actual: Weather Corrected	Actual	EB-2006-0501 Forecast	EB-2008-0272 Forecast	EB-2010-0002 Forecast	EB-2012 0031 Forecast	EB-2014-0140 Forecast	EB-2016-0160 Forecast
2005	17,702						17,701	18,355	0.01					
2006	17,401						17,419	18,031	-0.10					
2007	17,086	17,329					17,329	18,537	-1.40	0.00				
2008	17,142	17,386					17,413	17,611	-1.56	-0.16				
2009		17,376	17,333				17,333	16,999		0.25	0.00			
2010		16,905	16,999				16,839	17,551		0.39	0.95			
2011			16,850	16,769			16,769	17,274			0.48	0.00		
2012			16,667	16,718			16,645	17,292			0.14	0.44		
2013				16,759	16,606		16,606	17,536				0.92	0.00	
2014				16,400	16,748		16,819	17,007				-2.49	-0.42	
2015					17,060	16,731	16,731	16,952					1.96	0.00
2016					17,317	16,756	16,715	17,040					3.60	0.24
2017						16,872	16,306	16,247						3.47
2018						16,876	16,329	17,151						3.35
Average Excluding First Year (Actual) (7)									-1.02	0.16	0.52	-0.37	1.71	2.36

(1) Forecast: EB-2006-0501; Ex A; T14; S 3; P 19 of 20.
(2) Forecast: EB-2008-0272; Ex A; T14; S 3; P 22 of 24.
(3) Forecast: EB-2010-0002; Ex A; T14; S 3; P 19 of 21.
(4) Forecast: EB-2012-0031; Ex A; T15; S 2; P 22 of 24.
(5) Forecast: EB-2014-0140; Ex A; T15; S 2; P 20 of 23, settlement amount shown.
(6) Forecast: EB-2016-0160; Ex E1; T3; S 1; P 20 of 52.
(7) Compares actual-weather corrected with forecast (3 years of forecast for EB-2006-0501, EB-2008-0272, EB-2010-0002, EB-2012-0031, EB-2014-0140, and EB-2016-0160 forecast).

RATES FOR EXPORT TRANSMISSION SERVICE

1. INTRODUCTION

The Independent Electricity System Operator ("IESO") collects Export Transmission Service ("ETS") revenues and remits them on a monthly basis to Hydro One, whose transmission system is used to facilitate export transactions at the point of interconnection with the neighbouring markets.

2. EXPORT TRANSMISSION SERVICE TARIFF DESIGN

Since the initial setting of the ETS rate, there have been many competing views advanced by stakeholders with respect to the basis of the tariff design and appropriateness of the charge level. As a result, over the years, the ETS rate has been determined through a combination of stakeholder agreements and Board interim Decisions, informed by Board-directed studies performed by both the IESO, and most recently, by Hydro One Transmission.

As a part of Hydro One's 2015/2016 Transmission Rate Application (EB-2014-0140), Hydro One Transmission engaged Elenchus Research Associates ("Elenchus") to perform a cost allocation study of network assets utilized by export transmission customers to determine the ETS rate based on cost causality principles. The Elenchus study was stakeholdered with interested parties and a final report was included in Exhibit H1, Tab 5, Schedule 1, Attachment 1 of that application.

The criteria for Elenchus' recommended methodology to allocate costs are defined below:

- Utilize the prior year actual hourly data for domestic and export customers;

- Utilize the 12 Coincident Peak[1] ("CP") as the allocator in apportioning assets between domestic and export customers in order to develop composite allocators to allocate shared expenses;
- Allocate only dedicated assets used to serve export customers and related expenses to the export customer class. No asset related costs associated with shared assets should be allocated to export customers;
- Allocate OM&A expenses related to the use of shared assets to export customers using composite assets as allocator;
- Exclude external revenues from the allocation to the export customer class; and
- Calculate the ETS rate based on Hydro One Transmission's proposed Network revenue requirement, adjusted to include other transmitters' approved revenue requirement reflected in the Uniform Transmission Rates ("UTRs").

The cost allocation study completed by Elenchus recommended an ETS rate of $1.70/MWh for 2015 and 2016 as being reflective of the cost of providing export service.

For the purpose of reaching a settlement, all parties agreed to an ETS rate change from the $2.00/MWh, currently in effect at the time, to $1.85/MWh. This rate was approved by the Board in its EB-2014-0140 Decision as the effective rate for 2015 and 2016, and subsequently maintained as the effective rate for 2017 and 2018 in its EB-2016-0160 Decision.

In this application, Hydro One updated the 2015 Elenchus cost allocation model utilizing the latest available information. This included updates to: the fixed assets dedicated to interconnections, the 2018 system peak and export load data used to determine the 12 CP allocator, and the forecast for 2020 ETS exports (MWh). Based on the updated cost

[1] Domestic and Export Demand at Ontario system peak.

allocation model data and Hydro One's proposed 2020 revenue requirement, the 2020 ETS rate calculated using the Elenchus study methodology has been determined to be $1.25/MWh. The decrease in the calculated ETS rate as compared to the 2015 study primarily reflects a decrease in Hydro One's OM&A costs relative to what was proposed at the time the 2015 study was completed, and an increase in forecast exports (MWh) from what was assumed in the 2015 study. The following Table 1 demonstrates these key differences in the parameters utilized in 2015 Elenchus cost allocation study and the updated cost allocation study in this application.

Table 1: ETS Rates Derived Using Elenchus Cost Allocation Study

Year	Total Hydro One Revenue Requirement allocated to Export	ETS Exports (MWh)	ETS Rate ($/MWh)
2015	27.2 million	16,700,000	1.70
2020	22.1 million	18,800,000	1.25

While the updated cost allocation study resulted in a calculated ETS rate of $1.25/MWh, the current ETS rate of $1.85/MWh represents a negotiated rate that was established as part of the Settlement Agreement in Proceeding EB-2014-0410. In addition, a decrease in the ETS rate will negatively impact the transmission rates that Ontario customers pay and could be perceived as benefiting customers in neighbouring jurisdictions at the expense of Ontario consumers. As such, Hydro One proposes to continue using the current ETS rate of $1.85/MWh to establish the ETS revenue used to offset the transmission revenue requirement as discussed in Section 3.

3. EXPORT TRANSMISSION SERVICE REVENUE

Hydro One's ETS revenue, used for establishing the rates revenue requirement proposed in this Application, is calculated using the currently approved tariff of $1.85/MWh and the three year historical rolling average volume of electricity exported from, or wheeled-through, Ontario over its transmission system. Table 2 provides the forecast of ETS revenue for the period 2020 to 2022.

Table 2: ETS Revenue Forecast ($ Millions)

Year	ETS Revenue
2020	35.9
2021	35.9
2022	36.3

The ETS revenue will continue to be disbursed as a decrease to the revenue requirement for the Network rate pool, as per the cost allocation process approved by the Board.

Hydro One proposes to revise its rates revenue requirement to reflect the Board's Decision and Order with respect to the ETS tariff as part of the Draft Rate Order to be submitted in finalizing the 2020 Uniform Transmission Rates.

VECC INTERROGATORY #54

Reference:

I2-04-01 p. 1-3
EB-2014-0140 Decision
EB-2014-0140, HONI's Tx 2015-2016 Transmission Revenue Requirement Application – Application, Settlement Agreement and Evidence

Interrogatory:

a) Please confirm that the parties to the EB-2014-0140 agreed on the ETS rate on the understanding that the methodologies, assumptions and scenarios used in the Elenchus Study do not have precedential value and may be challenged in subsequent proceedings.

b) Please confirm that the Board, in its EB-2014-0140 Decision, did not opine on the merits of or accept the methodologies, assumptions and scenarios used in the Elenchus Study.

Response:

a) Confirmed. On page 25 of the Settlement Agreement in EB-2014-0140 it states that: "*agreement on the level of ETS rate of $1.85 per MWh shall not be construed as acceptance of the methodology, assumptions, or scenarios used in the Elenchus Study*" and further states that "*because this is the first case where a cost allocation study was filed in evidence to inform the ETS Rate, the parties observe that the cost allocation methodology proposed by the Elenchus Study remains untested and the parties do not necessarily agree with that methodology. The parties therefore agreed on the ETS rate on the understanding that the methodologies, assumptions and scenarios used in the Elenchus Study do not have precedential value and may be challenged in subsequent proceedings*."

b) Confirmed. In the OEB Decision recorded in the December 2, 2014 transcript of this proceeding, the OEB accepted and approved the Settlement Agreement as filed and did not opine on any matters specifically related to ETS or the Elenchus Study.

Witness: Clement Li, Henry Andre

UNDERTAKING - JT 1.36 - Q1

Reference:

I2-APPrO-1

I-10-VECC-55

Undertaking:

Preamble:

In response to I2-APPrO-1, Hydro One filed a copy of the 2015 Elenchus cost allocation model in live excel format with information that was updated to calculate the ETS Rate of $1.25/MWh.

Hydro One also provided the following table summarizing the calculation of the $1.25/MWh in response to I2-APPrO-1:

			UTR Network Revenue Requirement				
ETS Allocated Revenue Requirement ($M)	Volume (GWh)	Rate ($/MWh)	Hydro One Total ($M)	Ontario Total ($M)	Escalation Factor	Ontario ETS Revenue Requirement ($M)	Ontario ETS Rate ($/MWh)
A	B	C=A/B	D	E	F=E/D	G=A X F	H=G/B
$22.1	18,800.0	$1.17	$977.6	$1,041.9	106.6%	$23.5	$1.25

Note: All revenue requirement amounts are based on Hydro One's proposed 2020 revenue requirement, as shown in Exhibit I2, Tab 4, Schedule 1, Table 1.

In response to I-10-VECC-55, Hydro One explained that:

Response:

a) The export volumes for 2020 to 2022 were calculated based on a three year rolling average of the prior year's amounts. The table below provides the export volumes for 2020 to 2022 period as used in the initial Application:

2015 Actual	2016 Actual	2017 Actual	2018 (2015 - 2017 Avg)	2019 (2016 - 2018 Avg)	2020 (2017- 2019 Avg)	2021 (2018- 2020 Avg)	2022 (2019- 2021 Avg)
23,138,052	22,157,981	19,346,599	21,547,544	21,017,374	20,637,172	21,067,364	20,907,304

b) The same calculation as in part (a) was used for the Updated Application; however the data for 2018 was updated to reflect actual volumes. The table below provides the export volumes for 2020 to 2022 period as used in the Updated Application:

2015 Actual	2016 Actual	2017 Actual	2018 Actual	2019 (2016 - 2018 Avg)	2020 (2017- 2019 Avg)	2021 (2018- 2020 Avg)	2022 (2019- 2021 Avg)
23,138,052	22,157,981	19,346,599	18,771,464	20,092,015	19,403,359	19,422,279	19,639,218

a) In respect of I10-VECC-55, please explain the benefits of using a three-year rolling average to forecast export volumes.

Witness: Clement Li

b) Please confirm that Hydro One is forecasting 2020 export volumes in the Updated Application of 19,403,359 MWh, however Hydro One's calculation of the ETS Rate of $1.25/MWh assumes the allocated 2020 export revenue requirement of $22,080,665 is collected from an export volume of 18,800 GWh.

c) Please update the calculation of the ETS Rate assuming Hydro One's proposed 2020 export revenue requirement is collected from Hydro One's forecasted 2020 export volumes of 19,403,359 MWh. In connection with this update, please provide:
   a. the resulting ETS Rate,
   b. an update to the summary table that was provided in I2-APPrO-1 showing the values used for this scenario, and
   c. a revised version of the live excel version of the Elenchus cost allocation model updated to reflect this scenario.

d) Please update Hydro One's forecast of export volumes using a four-year rolling average methodology (rather than a three-year rolling average), and provide updated forecasts of export volumes for 2019, 2020, 2021, and 2022.

e) Please update the calculation of the ETS Rate assuming Hydro One's proposed 2020 export revenue requirement is collected from the forecast of 2020 export volumes calculated in response to part (d) above. In connection with this update, please provide:
   a. the resulting ETS Rate,
   b. an update to the summary table that was provided in I2-APPrO-1 showing the values used for this scenario, and
   c. a revised version of the live excel version of the Elenchus cost allocation model updated to reflect this scenario.

Response:

a) Normally, a three-year rolling average has the benefit that it captures the up and down fluctuations of prior years for a value being forecast. However, as shown in the response to Exhibit I, Tab 10, Schedule VECC-55 part (b), export volumes have been on a clear downward trend since 2015 and therefore the three-year average does not provide the best estimate of what the forecast exports will be in 2020.

b) Confirmed. For the purpose of cost allocation and rate design, Hydro One believes that an export volume of 18,800 GWh is the best estimate of export volumes in 2020.

c) The resulting ETS Rate assuming Hydro One's proposed 2020 export revenue requirement is collected from the forecasted 2020 export volumes using three-year rolling average of 19,403,359 MWh is $1.21/MWh.

Below is the updated table as requested:

			UTR Network Revenue Requirement				
ETS Allocated Revenue Requirement ($M)	**Volume (GWh)**	**Rate ($/MWh)**	**Hydro One Total ($M)**	**Ontario Total ($M)**	**Escalation Factor**	**Ontario ETS Revenue Requirement ($M)**	**Ontario ETS Rate ($/MWh)**
A	B	C=A/B	D	E	F=E/D	G=A X F	H=G/B
$22.1	19,403.4	$1.14	$977.6	$1,041.9	106.6%	$23.5	$1.21

A revised version of the live excel version of the Elenchus cost allocation model updated to reflect this scenario is provided as Attachment 1 to this undertaking.

d) The table below provides the requested information:

Export Volume Forecast using 4-year Rolling Average			
2019 Export MWh Forecast (2015 - 2018 Avg)	**2020 Export MWh Forecast (2016- 2019 Avg)**	**2021 Export MWh Forecast (2017- 2020 Avg)**	**2022 Export MWh Forecast (2018- 2021 Avg)**
20,853,524	20,282,392	19,813,495	19,930,219

e) The resulting ETS Rate assuming Hydro One's proposed 2020 export revenue requirement is collected from the forecasted 2020 export volumes using four-year rolling average of 20,282,392 MWh is $1.16/MWh.

Below is the updated table as requested:

			UTR Network Revenue Requirement				
ETS Allocated Revenue Requirement ($M)	**Volume (GWh)**	**Rate ($/MWh)**	**Hydro One Total ($M)**	**Ontario Total ($M)**	**Escalation Factor**	**Ontario ETS Revenue Requirement ($M)**	**Ontario ETS Rate ($/MWh)**
A	B	C=A/B	D	E	F=E/D	G=A X F	H=G/B
$22.1	20,282.4	$1.09	$977.6	$1,041.9	106.6%	$23.5	$1.16

A revised version of the live excel version of the Elenchus cost allocation model updated to reflect this scenario is provided as Attachment 2 to this undertaking.

2020 ONTARIO UNIFORM TRANSMISSION RATE SCHEDULES

EB-2019-xxxx

The rate schedules contained herein shall be effective January 1, 2020

Issued: Month, Year
Ontario Energy Board

TERMS AND CONDITIONS

(A) APPLICABILITY The rate schedules contained herein pertain to the transmission service applicable to: •The provision of Provincial Transmission Service (PTS) to the Transmission Customers who are defined as the entities that withdraw electricity directly from the transmission system in the province of Ontario. •The provision of Export Transmission Service (ETS) to electricity market participants that export electricity to points outside Ontario utilizing the transmission system in the province of Ontario. The Rate Schedule ETS applies to the wholesale market participants who utilize the Export Service in accordance with the Market Rules of the Ontario Electricity Market, referred to hereafter as Market Rules. These rate schedules do not apply to the distribution services provided by any distributors in Ontario, nor to the purchase of energy, hourly uplift, ancillary services or any other charges that may be applicable in electricity markets administered by the Independent Electricity System Operator (IESO) of Ontario.

(B) TRANSMISSION SYSTEM CODE The transmission service provided under these rate schedules is in accordance with the Transmission System Code (Code) issued by the Ontario Energy Board (OEB). The Code sets out the requirements, standards, terms and conditions of the transmitter's obligation to offer to connect to, and maintain the operation of, the transmission system. The Code also sets out the requirements, standards, terms and conditions under which a Transmission Customer may connect to, and remain connected to, the transmission system. The Code stipulates that a transmitter shall connect new customers, and continue to offer transmission services to existing customers, subject to a Connection Agreement between the customer and a transmitter.

(C) TRANSMISSION DELIVERY POINT The Transmission Delivery Point is defined as the transformation station, owned by a transmission company or by the Transmission Customer, which steps down the voltage from above 50 kV to below 50 kV and which connects the customer to the transmission system. The demand registered by two or more meters at any one delivery point shall be aggregated for the purpose of assessing transmission charges at that delivery point if the corresponding distribution feeders from that delivery point, or the plants taking power from that delivery point, are owned by the same entity within the meaning of Ontario's *Business Corporations Act*. The billing demand supplied from the transmission system shall be adjusted for losses, as appropriate, to the Transmission Point of Settlement, which shall be the high voltage side of the transformer that steps down the voltage from above 50 kV to below 50 kV.

(D) TRANSMISSION SERVICE POOLS The transmission facilities owned by the licenced transmission companies are categorized into three functional pools. The transmission lines that are used for the common benefit of all customers are categorized as Network Lines and the corresponding terminating facilities are Network Stations. These facilities make up the Network Pool. The transformation station facilities that step down the voltage from above 50 kV to below 50 kV are categorized as the Transformation Connection Pool. Other electrical facilities (i.e. that are neither Network nor Transformation) are categorized as the Line Connection Pool. All PTS customers incur charges based on the Network Service Rate (PTS-N) of Rate Schedule PTS. The PTS customers that utilize transformation connection assets owned by a licenced transmission company also incur charges based on the Transformation Connection Service Rate (PTS-T). The customer demand supplied from a transmission delivery point will not incur transformation connection service charges if a customer fully owns all transformation connection assets associated with that transmission delivery point. The PTS customers utilize lines owned by a licenced transmission company to connect to Network Station(s) also incur charges based on the Line Connection Service Rate (PTS- L). The customer demand supplied from a transmission delivery point will not incur line connection service charges if a customer fully owns all line connection assets connecting that delivery point to a Network Station. Similarly, the customer demand will not incur line connection service charges for demand at a transmission delivery point located at a Network Station.

(E) MARKET RULES The IESO will provide transmission service utilizing the facilities owned by the licenced transmission companies in Ontario in accordance with the Market Rules. The Market Rules and appropriate Market Manuals define the procedures and processes under which the transmission service is provided in real or operating time (on an hourly basis) as well as service billing and settlement processes for transmission service charges based on rate schedules contained herein.

(F) METERING REQUIREMENTS In accordance with Market Rules and the Transmission System Code, the transmission service charges payable by Transmission Customers shall be collected by the IESO. The IESO will utilize Registered Wholesale Meters and a Metering Registry in order to calculate the monthly transmission service charges payable by the Transmission Customers. Every Transmission Customer shall ensure that each metering installation in respect of which the customer has an obligation to pay transmission service charges arising from the Rate Schedule PTS shall satisfy the Wholesale Metering requirements and associated obligations specified in Chapter 6 of the Market Rules, including the appendices therein, whether or not the subject meter installation is required for settlement purposes in the IESO-administered energy market. A meter installation required for the settlement of charges in the IESO-administered that energy market may be used for the settlement of transmission service charges. The Transmission Customer shall provide to the IESO data required to maintain the information for the Registered Wholesale Meters and the Metering Registry pertaining to the metering installations with respect to which the Transmission Customers have an obligation to pay transmission charges in accordance with Rate Schedule PTS. The Metering Registry for metering installations required for the calculation of transmission charges shall be maintained in accordance with Chapter 6 of the Market Rules. The Transmission Customers, or Transmission Customer Agents if designated by the Transmission Customers, associated with each Transmission Delivery Point will be identified as Metered Market Participants within the IESO's Metering Registry. The metering data recorded in the Metering Registry shall be used as the basis for the calculation of transmission charges on the settlement statement for the Transmission Customers identified as the Metered Market Participants for each Transmission Delivery Point. The Metering Registry for metering installations required for calculation of transmission charges shall also indicate whether or not the demand associated with specific Transmission Delivery Point(s) to which a Transmission Customer is connected attracts Line and/or Transformation Connection Service Charges. This information shall be consistent with the Connection Agreement between the Transmission Customer and the licenced Transmission Company that connects the customer to the IESO-Controlled Grid.

(G) EMBEDDED GENERATION The Transmission Customers shall ensure conformance of Registered Wholesale Meters in accordance with Chapter 6 of Market Rules, including Metering Registry obligations, with respect to metering installations for embedded generation that is located behind the metering installation that measures the net demand taken from the transmission system if (a) the required approvals for such ~~generation~~ generator unit or energy storage facility are obtained after October 30, 1998; and (b) the generator unit nameplate rating is 2 MW or higher for renewable generation and 1 MW or higher for non- renewable generation or if the individual inverter unit capacity is 1 MW or higher for energy storage or solar generators; and (c) the Transmission Delivery Point through which the generator or energy storage facility is connected to the transmission system attracts Line or Transformation Connection Service charges. These terms and conditions also apply to the incremental capacity associated with any refurbishments or expansions approved after October 30, 1998, to a generator or generation facility ~~unit~~ that was connected through an eligible Transmission Delivery Point on or prior to October 30, 1998 and the approved incremental generator nameplate capacity is 2 MW or higher for renewable generation and 1 MW or higher for non-renewable generation or if the individual inverter unit capacity is 1 MW or higher for expansion of energy storage facilities or solar generators. The term renewable generation refers to a facility that generates electricity from the following sources: wind, solar, Biomass, Bio-oil, Bio-gas, landfill gas, or water. Accordingly, the distributors that are Transmission Customers shall ensure that connection agreements between them and the generators, load customers, and embedded distributors connected to their distribution system have provisions requiring the Transmission Customer to satisfy the requirements for Registered Wholesale Meters and Metering Registry for such embedded generation even if the subject embedded generator(s) do not participate in the IESO-administered energy markets.

(H) EMBEDDED CONNECTION POINT In accordance with Chapter 6 of the Market Rules, the IESO may permit a Metered Market Participant, as defined in the Market Rules, to register a metering installation that is located at the embedded connection point for the purpose of recording transactions in the IESO-administered markets. (The Market Rules define an embedded connection point as a point of connection between load or generation facility and distribution system). In special situations, a metering installation at the embedded connection point that is used to settle energy market charges may also be used to settle transmission service charges, if there is no metering installation at the point of connection of a distribution feeder to the Transmission Delivery Point. In above situations: •The Transmission Customer may utilize the metering installation at the embedded connection point, including all embedded generation and load connected to that point, to satisfy the requirements described in Section (F) above provided that the same metering installation is also used to satisfy the requirement for energy transactions in the IESO- administered market. •The Transmission Customer shall provide the Metering Registry information for the metering installation at the embedded connection point, including all embedded generation and load connected to that point, in accordance with the requirements described in Section (F) above so that the IESO can calculate the monthly transmission service charges payable by the Transmission Customer.

RATE SCHEDULE: (PTS) PROVINCIAL TRANSMISSION RATES

APPLICABILITY:

The Provincial Transmission Service (PTS) is applicable to all Transmission Customers in Ontario who own facilities that are directly connected to the transmission system in Ontario and that withdraw electricity from this system.

	Monthly Rate ($ per kW)
Network Service Rate (PTS-N):	**4.35**
$ Per kW of Network Billing Demand [1,2]	
Line Connection Service Rate (PTS-L):	**0.83**
$ Per kW of Line Connection Billing Demand [1,3]	
Transformation Connection Service Rate (PTS-T):	**2.44**
$ Per kW of Transformation Connection Billing Demand [1,3,4]	

The rates quoted above shall be subject to adjustments with the approval of the Ontario Energy Board.

Notes:
1 The demand (MW) for the purpose of this rate schedule is measured as the energy consumed during the clock hour, on a "Per Transmission Delivery Point" basis. The billing demand supplied from the transmission system shall be adjusted for losses, as appropriate, to the Transmission Point of Settlement, which shall be the high voltage side of the transformer that steps down the voltage from above 50 kV to below 50 kV at the Transmission Delivery Point.

2. The Network Service Billing Demand is defined as the higher of (a) customer coincident peak demand (MW) in the hour of the month when the total hourly demand of all PTS customers is highest for the month, and (b) 85 % of the customer peak demand in any hour during the peak period 7 AM to 7 PM (local time) on weekdays, excluding the holidays as defined by IESO. The peak period hours will be between 0700 hours to 1900 hours Eastern Standard Time during winter (i.e. during standard time) and 0600 hours to 1800 hours Eastern Standard Time during summer (i.e. during daylight savings time), in conformance with the meter time standard used by the IMO settlement systems.

3. The Billing Demand for Line and Transformation Connection Services is defined as the Non-Coincident Peak demand (MW) in any hour of the month. The customer demand in any hour is the sum of (a) the loss-adjusted demand supplied from the transmission system plus (b) the demand that is supplied by an embedded generator unit or energy storage facility for which the required government approvals are obtained after October 30, 1998 and which have installed nameplate capacity of 2MW or more for renewable generation and 1 MW or higher for non-renewable generation or if the individual inverter unit capacity is 1 MW or higher for energy storage or solar generators, ~~on~~ or the demand supplied by the incremental capacity associated with a refurbishment or expansion approved after October 30, 1998, to a generator ~~unit~~ or generation facility that existed on or prior to October 30, 1998. The term renewable generation refers to a facility that generates electricity from the following sources: wind, solar, Biomass, Bio-oil, Bio-gas, landfill gas, or water. The demand supplied by embedded generation will not be adjusted for losses.

4. The Transformation Connection rate includes recovery for OEB approved Low Voltage Switchgear compensation for Toronto Hydro Electric System Limited and Hydro Ottawa Limited.

TERMS AND CONDITIONS OF SERVICE:

The attached Terms and Conditions pertaining to the Transmission Rate Schedules, the relevant provisions of the Transmission System Code, in particular the Connection Agreement as per Appendix 1 of the Transmission System Code, and the Market Rules for the Ontario Electricity Market shall apply, as contemplated therein, to services provided under this Rate Schedule.

RATE SCHEDULE: (ETS) EXPORT TRANSMISSION SERVICE

APPLICABILITY:

The Export Transmission Service is applicable for the use of the transmission system in Ontario to deliver electrical energy to locations external to the Province of Ontario, irrespective of whether this energy is supplied from generating sources within or outside Ontario.

	Hourly Rate
Export Transmission Service Rate (ETS):	$1.85 / MWh

The ETS rate shall be applied to the export transactions in the Interchange Schedule Data as per the Market Rules for Ontario's Electricity Market. The ETS rate shall be subject to adjustments with the approval of the Ontario Energy Board.

TERMS AND CONDITIONS OF SERVICE:

The attached Terms and Conditions pertaining to the Transmission Rate Schedules, the relevant provisions of the Transmission System Code and the Market Rules for the Ontario Electricity Market shall apply, as contemplated therein, to service provided under this Rate Schedule.

UNDERTAKING - JT 2.34 - Q18

Reference:
Exhibit I/Tab 01/Schedule 225 b) (OEB Staff-225 b))
Exhibit I2/Tab 6/Schedule 1, Attachment 1, page 3

Undertaking:
Preamble: The response to OEB Staff 225 b) states: "It is Hydro One's interpretation and practice to include customers with energy storage facilities and/or solar generators (the individual inverter with capacity is 1 MW or higher) in the data provided to the IESO for billing Line Connection and Transformation Connection customers on a gross load basis as per the approved UTR tariff".

It is noted that in the currently approved 2019 Uniform Transmission rates, renewable embedded generation only attracts Line and Transformation Connection charges if the generator unit rating is 2 MW or greater and the 1 MW cut-off applies to non-renewable generators.

a) Please explain why the cut-off for energy storage and solar generators is 1 MW and not 2 MW, particularly in the case of solar generators.

Response:
a) Energy storage is not considered renewable generation and therefore the cut-off is 1 MW.

Hydro One's experience is that solar inverter unit capacity is typically small (i.e. under 0.5 MW) and therefore the 1 MW limit is irrelevant. In any case, currently no Hydro One transmission customers with embedded solar generation are billed on a gross load basis and therefore this condition is not applied.

Hydro One agrees that solar generators are renewable generation and therefore the cut-off should be 2 MW. As such, Hydro One proposes to remove the words "or solar generators" from the following exhibits:

- Exhibit I2, Tab 2, Schedule 1, page 4, lines 13 and 18;
- Exhibit I2, Tab 2, Schedule 1, page 5, lines 15 and 20;
- Exhibit I2, Tab 6, Schedule 2, Attachment 1, page 3 Terms and Conditions (G); and

- Exhibit I2, Tab 6, Schedule 2, Attachment 1, page 5 Notes 3.

OEB INTERROGATORY #225

Reference:
I2-02-01 p.3 of 5

Interrogatory:
At the reference above, it is stated that:

Hydro One is proposing to update the definition of billing demand for the Line and Transformation Connection services to reflect the changes in the embedded generation market over the years, such as inclusion of energy storage facilities.

The "Embedded Generation" section in the proposed 2020 Ontario Uniform Transmission Rate Schedules (Exhibit I2, Tab 6, Schedule 2, Attachment 1) has also been updated to align with the changes in billing demand for the Line and Transformation Connection services.

a) Please explain why the proposed changes to the definition of billing demand for the line and transformation connection services and the changes to the definition of embedded generation are necessary.

b) Please discuss whether or not there are costs shifted to other customers if existing customers with energy storage facilities and/or solar generators (the individual inverter unit capacity is one MW or higher) are continuing to be billed on a net load basis.
    i. If so, please quantify the shifted costs.
    ii. If not, why not.

c) Please explain when and how the original definitions were determined.

d) Did Hydro One consult customers with energy storage facilities and/or solar generators (the individual inverter unit capacity is 1 MW or higher) about the proposed changes? If so, what are the customers' feedback on the proposed changes? If not, why not?

e) Please estimate the bill impact for a customer with energy storage facilities or solar generators before and after the proposed changes using the proposed 2020 UTRs.

Response:

a) The definition of billing demand for the line and transformation connection services and embedded generation in the current Uniform Transmission Rate ("UTR") Schedules have not been updated since 2005[1]. The proposed changes in wording clarify and reflect Hydro One's interpretation of these definitions in the data provided to the IESO for transmission billing purposes.

b) It is Hydro One's interpretation and practice to include customers with energy storage facilities and/or solar generators (the individual inverter with capacity is 1 MW or higher) in the data provided to the IESO for billing Line Connection and Transformation Connection customers on a gross load basis as per the approved UTR tariff. As discussed, in part (a), the proposed wording changes simply clarify and reflect Hydro One's interpretation. There will be no cost shifting as there will be no change in Hydro One's practice.

c) The original definitions were approved in the OEB's May 26, 2000 Decision on Hydro One's Transmission Application (RP-1999-0044). Section 3.2 of this Decision provides the rationale.

A 2 MW limit for renewable generation was added and was approved by the OEB in the Transmission System Code Phase 1 Policy Decision with Reasons (RP-2002-0120, issued June 8, 2004). Section 5.2 of this Decision provides the rationale. Subsequently, the UTR Schedule was updated under EB-2005-0241.

d) Hydro One did not consult customers about the proposed changes in wording. As discussed in parts (a) and (b), these wording changes simply clarify and reflect Hydro One's interpretation and practice. Customers will not be impacted by these changes.

e) Does not apply. Please see response in part (b).

[1] OEB Decision and Order EB-2005-0241 issued December 8, 2005

Power System Engineering, Inc.

Full-service consultants

Transmission Study for Hydro One Networks:

Recommended CIR Parameters and Productivity Comparisons

Prepared by:

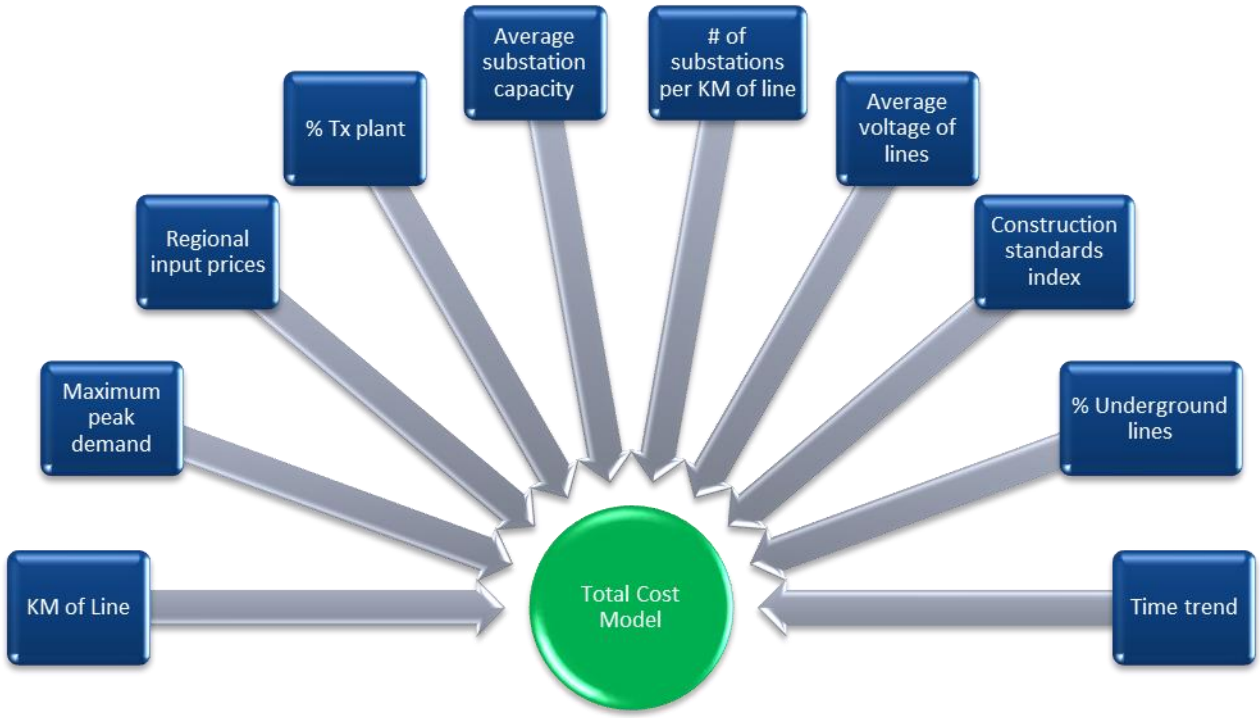
Power System Engineering, Inc.

January 24, 2019

List of Tables

List of Figures

Figure 1 Total Cost Model Variables



The benchmark scores are derived by taking the logarithmic percentage difference between Hydro One's actual total costs and their model-predicted total costs. A negative number implies that the company's actual costs are lower than the benchmark (i.e., lower than expected for an average utility with that company's operating circumstances). Table 2 and Figure 2 show Hydro One's scores for the historical and projected years.

LPMA INTERROGATORY #5

Reference:

A-04-01-01

Interrogatory:

a) The report is dated January 24, 2019. Please explain why data beyond 2017 was not used.

b) Please updates Tables 2 and 3 in the report to reflect actual data for 2017 and 2018. Please explain fully if this cannot be done for any of the years requested.

Response:

a) Please see the response to I-01-OEB 8(b).

b) Please see the response to I-01-OEB 8(b).

Witness: Stephen Vetsis, Steve Fenrick

OEB INTERROGATORY #8

Reference:

A-04-01-01 p.5,7,17 & 21
EB-2018-0218, OEB Staff IR 59 (Exhibit I/Tab 1/Schedule 59)
EB-2018-0218 Revised Public Redacted Technical Conference Transcript Volume 2 (January 15, 2019), p. 39/l. 6 to p. 162/l. 14
Decision and Order EB-2018-0218, p. 21
EB-2018-0218 Exhibit I/Tab 1/Schedule 65 (OEB Staff IR # 65)
EB-2018-0218 JT2.8
EB-2018-0218 JT2.9

Interrogatory:

On page 5 of its evidence, PSE notes the following changes have been made to its methodology in its evidence from that filed in the Hydro One Sault Ste. Marie revenue cap application:

This report has been revised from the Power System Engineering, Inc. (PSE) report filed in the Hydro One Sault Ste. Marie LP (SSM) application found in EB-2018-0218. Our recommendations regarding the customer incentive regulation parameters remain unchanged and our findings are similar to the report previously filed. No changes to the study have been made except the modifications which are listed and explained below.

1. Hydro One Networks provided PSE with a revised business plan that includes modified OM&A and capital spending levels for the projected years of the study.

2. A second modification has occurred due to PSE identifying peak demand data that was incorrectly reported by the three Southern Companies (Alabama Power, Gulf Power, and Mississippi Power) included in the sample. This data has now been corrected.[1]

3. The third modification are slight revisions in plant additions in 2016 and 2017 made by Hydro One.

Witness: Steve Fenrick

4. The incentive regulation period moves to 2020 to 2022 which means the OM&A spending is now escalated for 2021 and 2022 by I-X using the 2020 test year expenses rather than 2019.

5. Two minor corrections in the code were made relative to the prior research. The first is we are now calculating the asset prices prior to 1963 in calculating the capital benchmarks. The second is including only the observations in the TFP sample when aggregating the TFP components.[2]
   These five modifications have been incorporated into this revision and are the only changes made to the dataset and study methodology relative to the research filed EB-2018-0218 and EB-2018-0130.

6. This adjustment moved the TFP annual trend upwards by around 0.42%.

7. Both corrections had a minimal impact on the results with the effect of the change being a slightly lower TFP trend by around 0.16%.

PSE notes on page 17 of its evidence that the long-term TFP trend is -1.45%, a change of -0.16 percentage points from the study filed in EB-2018-0218. The historical data range for the TFP analysis was unchanged, from 2004 to 2016, as noted on page 7 of PSE's evidence.

a) Please confirm that only the changes in bullets 2, 3 (with respect to 2016 capital additions for Hydro One), and 5 relate to PSE's TFP analysis.

b) Please explain why PSE did not update its TFP and total cost benchmarking analysis with an additional year of data of 2017 actuals for both Hydro One Networks and the U.S. sample.

c) In its analyses documented in its evidence filed in this Application and in Hydro One SSM's application in EB-2018-0218, PSE has introduced a new constructed variable which it terms as a "loading" or "engineering construction index" to measure regional standards for the physical construction of networks to withstand climactic extremes for wind speeds, storms, ice loading, etc. OEB staff have used the term "hardening" as "loading" can also be used in the context of capacity or over-loading of electrical equipment.

In the EB-2018-0218 Decision and Order with respect to Hydro One SSM's revenue cap plan, the OEB stated:

> The OEB reserves judgement on the new "construction standards index" variable[57] provided in the PSE Report. This new variable is worthy of further consideration, yet the concept was not fully vetted in this proceeding. Further, the OEB questions its relevance to Hydro One SSM and its asset base.
>
> [57] PSE defines this new variable in Exhibit D, Tab 1, Schedule 1, pp. 25-26 and Appendix A. OEB staff used the term "hardening" variable, as it refers to the engineering standard to which network infrastructure must be constructed to withstand climactic conditions, such as wind and ice, in different regions (OEB staff submission, April 12, 2019, pp. 28-29).

OEB staff recognizes that Decision and Order EB-2018-0218 was issued on June 20. 2019, after the filing of Hydro One's current Application. However, there was testing of this new variable during the EB-2018-0218 case, through interrogatories and during the Technical Conference. In particular, OEB staff raised a concern regarding the construction of the variable for Hydro One in that, based on Platts' GIS mapping, where all, or nearly all, of Ontario is used, while Hydro One has few or no transmission assets in a large portion of northern Ontario.

d) Please explain why PSE (or Hydro One) did not update the "hardening" variable in light of the record in EB-2018-0218.

e) Please confirm that item 2 of the changes noted on page 5 of the updated PSE report correspond to the problem identified in OEB Staff Interrogatory #65 from the SSM proceeding.

f) Please confirm that item 5 of the changes noted on page 5 correspond to the problems identified in undertakings JT2.8 (aggregation) and JT2.9 (asset price) from the Hydro One SSM technical conference.

<u>Response:</u>

a) Confirmed. The changes in bullets 2, 3 (with respect to 2016 capital additions for Hydro One), and 5 relate to PSE's Industry TFP analysis. Note, this does not apply to the Hydro One TFP analysis through to 2022.

b) When PSE prepared its report for this proceeding, and given the nature of the report (as a follow-up to the report in EB-2018-0218), it did not seem necessary to update it for the 2017 actual data for the sample utilities at that time. In EB-2018-0218 both the PSE and PEG studies and reports used 2016 as the most recent year of actual data for the sample utilities. The study could be updated for the more recent actual data, as necessary. PSE would be prepared to do so, but significant additional time would be required given the scope of work involved.

c) PSE did examine the transmission loading variable to see what the impact on the variable value may be if it was calculated based on the location of Hydro One's transmission lines rather than service territory. We found that the variable value for Hydro One would increase if the methodology were modified. In other words, the cost challenges to Hydro One due to the loading variable would increase if we modified the methodology. This likely would raise Hydro One's total cost benchmark and improve the benchmarking score.

d) However, we did not believe it was appropriate to modify the variable from what was used in EB-2018-0218 because of consistency concerns with the rest of the sample. Since we cannot institute a change from service territory mapping to transmission line mapping for the rest of the sample, we chose to continue to be consistent in calculating the variable between Hydro One and the sampled utilities using the designated service territory mappings as the basis for the variable.

e) Confirmed. Please note PSE's response in part b of OEB Staff Interrogatory #65.

f) Confirmed.

Hydro One Networks Inc.
7th Floor, South Tower
483 Bay Street
Toronto, Ontario M5G 2P5
www.HydroOne.com

Tel: (416) 345-5680
Cell: (416) 568-5534
frank.dandrea@HydroOne.com

Frank D'Andrea
Vice President, Regulatory Affairs & Chief Risk Officer

BY RESS, EMAIL AND COURIER

October 15, 2019

Ms. Kirsten Walli
Board Secretary
Ontario Energy Board
Suite 2700, 2300 Yonge Street
P.O. Box 2319
Toronto, ON M4P 1E4

Dear Ms. Walli,

EB-2019-0082 - Hydro One Networks Inc.'s 2020-2022 Transmission Custom IR Application (the "Application") - Reply Report regarding Pacific Economics Group report and interrogatory responses

Further to the correspondence of Hydro One's counsel dated October 10, 2019 which explained that given the various new issues and points raised in the Pacific Economics Group report and accompanying IR responses, Hydro One intended to file a reply report from its consultant, please find the report enclosed.

This filing has been submitted electronically using the OEB's Regulatory Electronic Submission System and two (2) hard copies will be sent via courier.

Sincerely,

ORIGINAL SIGNED BY FRANK D'ANDREA

Frank D'Andrea
cc. EB-2019-0082 parties (electronic)

Power System Engineering, Inc.

Full-service consultants

Reply to
PEG's Report ("Incentive Regulation for Hydro One Transmission")

Prepared by:

Power System Engineering, Inc.
October 15, 2019

Reply to
PEG's Report ("Incentive Regulation for Hydro One Transmission")

Contact

Steve Fenrick 608.334.5994

Table of Contents

1 Overview and Research Results

This report ("PSE Reply Report") is in reply to the report of PEG dated September 5, 2019 ("PEG Report") and the accompanying recent IR responses (the final ones delivered on October 9, 2019) from PEG which raise several new issues and points.

In response to specific concerns raised in the PEG Report regarding the length of the sample period, we are now able to provide in this report a two-year update to PSE's research found in the PSE Report, bringing the last year of the sample from 2016 to 2018.[1] In reply to the PEG Report we also point out two flaws in PEG's research and respond to other issues raised in the PEG Report.

In the PEG Report, PEG has produced benchmarking results for Hydro One that are not consistent with PEG's own results found in the Hydro One Sault Ste. Marie LP ("HOSSM") application.[2] In the HOSSM proceeding, PSE's results indicated Hydro One Networks should have a stretch factor of 0.0%, and PEG's corrected results indicated Hydro One Networks should have a stretch factor of 0.15%.[3] We discuss in this PSE Reply Report why PEG's results in this Hydro One application differ significantly from both PSE's results and PEG's own results in the HOSSM application.

Both PSE and PEG find negative total factor productivity ("TFP") trends in the electric transmission industry. Both consultants' TFP results are quite similar when examined over the same time periods. Where one disagreement lies is regarding the most appropriate time period to apply to Hydro One's Custom IR plan.

PEG is recommending a -0.25% productivity factor (we estimate this would become -0.44% if PEG added 2017 and 2018 data to its analysis). PSE finds a -1.61% industry TFP trend from 2004 to 2018 and is recommending a 0.0% productivity factor, because the Board has stated that it does not wish to have negative productivity factors. PEG recommends a total X-factor of 0.05% (-0.25% productivity factor plus 0.3% stretch factor). PSE is recommending a 0.0% X-factor (0.0% productivity factor plus 0.0% stretch factor). We note that if the productivity factor is set to 0.0% by the Board, both studies show there is an implicit stretch factor already embedded in the productivity factor. Further, Hydro One's proposal already includes a progressive productivity proposal that amounts to a 0.14% stretch factor in 2021 and a 0.33% stretch factor in 2022.[4]

1.1 Total Cost Benchmarking Results

In the HOSSM application, PEG corrected some of its errors identified by PSE in their response found in EB-2018-0218, Exhibit L1, Tab 1, Schedule 6. When PEG corrected these errors, Hydro One's benchmark score for 2020 to 2022 became -11.0%. In PEG's benchmarking research in this

---

[1] PEG did not update its report in this case. In PEG's reply to EB-2019-0082, L1, Tab 1, Schedule 6, part c PEG states that it was unable to provide an update to 2018 due to a lack of time.

[2] EB-2018-0218

[3] In the HOSSM case, PEG's benchmarking dataset had several errors in the old capital data that was used in PEG's report. When PEG corrected those errors, their results found in EB-2018-0218, Exhibit L1, Tab 1, Schedule 6, part i indicated Hydro One was -12% below benchmark costs, indicating a stretch factor of 0.15%.

[4] EB-2019-0082, JT 2.42.

application, PEG also corrected these errors.[5] In Table 1 below, we report PEG's benchmarking results in the HOSSM case using their results after PEG made the corrections in L1, Tab 1, Schedule 6 (part i).

PSE's results in the current application have been updated to include 2017 and 2018 actual data for the sample, and 2018 actual data for Hydro One. This is in response to PEG's comments regarding the length of the sample period for our research.[6] The 2017-18 data also provides the Board with the most recently available information. No methodological changes, other than updating data to 2018, have been implemented in this PSE Reply Report relative to the PSE Report.

The following table provides the benchmark scores of PSE and PEG in the HOSSM and Hydro One applications. The first column with results in green provides PSE's 2018 updated benchmark scores. The next two columns provide the results we presented in the PSE Report and then in the HOSSM application. The PEG results from the HOSSM application and the current application are shown in the last two columns.

Table 1 Total Cost Benchmarking Results of PSE and PEG

Year	PSE (Current Application with 2018 Sample Update)	PSE (Current Application but Sample only to 2016)	PSE (HOSSM, Sample only to 2016)	PEG (HOSSM, Sample only to 2016)[7]	PEG (Current Application with Sample only to 2016)
Average 2004-2018	-26.0%	-20.7%	-25.7%	-31.2%	-11.4%
Average 2016-2018	-29.5%	-24.4%	-30.2%	-20.4%	+1.0%
Average 2020-2022	-32.9%	-27.1%	-31.8%	-11.0%	+9.0%

PEG's HOSSM research indicated a 0.15% stretch factor, now PEG's results indicate a 0.3% stretch factor recommendation. PEG's 2020-2022 average score changed by 20% (from -11% to +9%) over a six-month time, despite PEG using the same sample period and benchmarking the same company. We anticipated PEG's results would move in the opposite direction, due to: (1) the company revising its business plan spending to lower levels relative to what was inputted in HOSSM, and (2) PEG endeavoring to exclude certain cost categories for Hydro One to make the cost definitions consistent (which PEG had not done in the HOSSM research).[8]

[5] See p. 59 in the PEG Report for a list of methodological changes. The first bullet point states that PEG has made the corrections from the errors identified in L1, Tab 1, Schedule 6 of the HOSSM proceeding.

[6] See 19 and 22 of the PEG Report.

[7] From EB-2018-0218, Exhibit L1, Tab 1, Schedule 6, part i (b).

[8] PEG did not actually subtract these costs during the forecasted years, and improperly subtracted them in the years of

Having reviewed in detail the PEG Report and accompanying working papers and IR responses, our view is that there are two main reasons – which in our view are flaws in PEG's research – why PEG's benchmark results for Hydro One have now changed so dramatically and do not align with PSE's analysis.

1.1.1 PEG's Model Is Biased Against Recent and Forecasted Time Periods

PEG's results indicate a rapid increase of the cost benchmarking scores for Hydro One. PEG's 2004-2018 average score for Hydro One is -11.4%, however by 2020-2022 the score has risen to +9.0%. In contrast, PSE's results demonstrate a moderate decline (i.e., improvement) in Hydro One's benchmark scores over time. The reason for this difference is that PEG's model contains a clear bias against the recent and forecasted years. This bias is against all the sampled utilities, including Hydro One. When this bias is resolved, PEG's results for Hydro One's 2020-2022 period change considerably.

In Section 2.1, we will demonstrate the clear bias in PEG's results against the recent and forecasted years. The PEG bias unfairly raises all the benchmark scores for all utilities during the recent years of the sample. By 2018, we estimate the bias to be substantial (+15%) and growing. That is to say, the entire sample's average benchmark score in 2018 is +15% rather than the expected 0%.[9] This bias should not exist in a properly specified model.

The bias in PEG's model can be resolved by inserting one variable (a quadratic trend variable) into the model. PEG should agree with the insertion of this variable because PEG itself has stated that this variable would be of interest to capture a curvature of costs which is what is seen in PEG's model.[10] When this variable is inserted and making no other changes, PEG's results would indicate that Hydro One's 2020 to 2022 benchmark score would improve by 25.1%. This would indicate a 0.15% stretch factor.

1.1.2 PEG Has Introduced a Different and Needless Modeling Procedure

The principal reason for the considerable change in PEG's scores is that PEG instituted a different modeling procedure that PEG did not use approximately six months ago in HOSSM, or in Hydro One's Distribution application from last year.[11] The different modeling procedure affects the

---

2008 to 2017. PSE noticed the error and PEG acknowledged it and provided corrected results in L1, Tab 1, Schedule 21 (a) and (b). The Hydro One 2020 to 2022 average benchmark score moved from +9.0% to +6.8% due to this correction.

[9] We would expect 0% to be the average benchmark score for the sample because this would indicate the average utility is at their benchmark (or expected) total costs. The objective of performance benchmarking is to provide a comparative analysis showing how a utility's costs compare to a hypothetical average utility sharing the same characteristics as that utility. PEG's model is not producing those results but, instead, is calculating the benchmarks to be 15% lower in 2018 than what an average utility would be expected to achieve.

[10] See Section 2 for Dr. Lowry's quote on the merits of including a quadratic trend variable in the most recent Toronto Hydro proceeding.

[11] PEG stated in EB-2017-0049, Exhibit L1, Tab 8, Schedule HONI-53 that PEG did not conduct an autocorrelation adjustment in its research of Hydro One Distribution. This statement is contradicted by PEG's statement in this case in EB-2019-0082, L1, Tab 1, Schedule 24, part c and d where PEG claims it did conduct an autocorrelation adjustment

underlying data that enters the regression and needlessly impacts the benchmarks, when more modern and standard procedures are available that do not influence the results. As we will discuss in Section 2, we believe that PEG's approach:

(1) Introduces a possibility for error, given the complex coding necessary to undertake its new modeling procedure,

(2) Has not proven to be a valid procedure on an unbalanced panel dataset,

(3) Is not necessary since the coefficients from the standard OLS run cannot be improved upon,

(4) Is open to subjective judgement by the researcher, and

(5) Is not easily reproduced and verified by non-experts using standard econometric software packages.[12]

Given the complexity and customization of PEG's econometric coding, we are unable to verify that PEG's new modeling procedure is being calculated properly. A more transparent and reproducible method would be to use commercially available econometric software packages that could easily reproduce PEG's results. We are forced to assume without verification that PEG conducted all of its coding properly, that these procedures are valid ones to implement, and that PEG made reasonable assumptions on the underlying sources of heteroskedasticity and autocorrelation when making these complex and unnecessary adjustments.[13] This assumption is made more difficult when PEG's results change considerably from six months prior.

If PEG's modeling procedure used the more modern procedure (or simply reverted to what PEG did in the HOSSM study) with no other changes made, PEG's results would indicate a 0.15% stretch factor for Hydro One as Hydro One's total cost benchmark score would be -20.5%.

If both these flaws are corrected— i.e., (1) resolve PEG's bias against the sample in recent years, and (2) switch to the modeling procedure to use the OLS coefficients or at the very least, the modeling procedure used by PEG in HOSSM and the Hydro One Distribution application—PEG's benchmarking scores for Hydro One would indicate a 0.0% stretch factor and show results consistent with PSE's analysis.

in its Hydro One Distribution research.

[12] While PEG's modeling procedure requires extensive customized code to be written with little ability to identify errors, PSE's benchmarks can be replicated by most off-the-shelf econometric software packages. In fact, in our working papers we provided results from two such vendors (EViews and STATA). The procedures have been vetted by thousands of users.

[13] PSE's modern approach requires no assumptions on the underlying sources of heteroskedasticity and autocorrelation, taking this subjective task out of the hands of the researcher.

1.2 Electric Transmission Industry Productivity Results

The second key component after benchmarking is used to set the stretch factor is calculating the industry TFP trend to formulate the productivity factor.[14] The PSE and PEG TFP results are quite similar over the same sample period. The difference in the productivity results are a consequence of different time periods employed. Over the sample period of 2005 to 2016, PSE calculates an industry TFP trend of -1.45%, and PEG calculates an industry TFP trend of -1.47%.[15] Similarly, if we examine the industry trend after 2010, PSE calculates an industry 2011 to 2016 TFP trend of -2.39%, and PEG calculates an industry TFP trend of -2.33% over the same time period.

In response to PEG and stakeholder comments and questions on lengthening the time period, PSE has now added the years 2017 and 2018 to our industry TFP sample. The sample starts in 2005 and goes to 2018. This provides 14 sampled years of TFP trends and incorporates the most recently available data. In conducting this update, we did not make any other changes to our methodology other than adding 2017 and 2018 observations to the industry sample.

The table below provides the industry TFP growth rates of both PSE and PEG for the current application and the HOSSM application.[16]

Table 2 PSE and PEG TFP Growth Rates

Year	PSE TFP Growth Rates (Current Application with 2018 Update)	PSE TFP Growth Rates (HOSSM, sample only goes to 2016)	PEG TFP Growth Rates (Current Application, sample only goes to 2016)	PEG TFP Growth Rates (HOSSM, sample only goes to 2016)[17]
1996-2016			-0.25%	-0.36%
2005-2016	-1.45%	-1.71%	-1.47%	-1.88%
2005-2018	**-1.61%**			
2017 and 2018	-2.42%			

The 2017 and 2018 years have continued the recent strongly negative decline in industry TFP. The 2017 and 2018 results show that using the more recent sample period of 2005-2016 is a far better predictor of the 2017 and 2018 TFP trends than the less applicable time period of 1996 to 2016. PSE's opinion is that the more contemporary period of 2005-2018 will continue to be the better predictor of the upcoming TFP trends in 2021 and 2022.

---

[14] PEG refers to the productivity trend as a multifactor productivity trend (MFP). We use the term TFP in the PSE Report and this Reply Report.

[15] In PEG's HOSSM research PEG found a TFP trend during this sample period of -1.82%.

[16] We again show the results produced by PEG's interrogatory response in HOSSM where PEG fixed their capital data. This was in EB-2018-0218, Exhibit L1, Tab 1, Schedule 6, part i (c). Their TFP results in that response did not have the large change that their benchmarking results had.

[17] EB-2018-0218, Exhibit L1, Tab 1, Schedule 6, part i (c)

Given the similarity in results, PSE would anticipate that if PEG updated its sample period to 2018, PEG's TFP estimate over their full sample period starting in 1996 would decline from their estimated -0.25% average trend. If we use the same TFP trends calculated by PSE for 2017 and 2018, PSE estimates the PEG average TFP trend for the 1996 to 2018 time period would become -0.44%. If PEG continued to base its X-factor recommendation on the industry productivity trend, this would lower their productivity factor recommendation to -0.44% and result in a negative X-factor recommendation.

2 Flaws in PEG's Benchmarking Research

The September 2019 PEG Report contains several flaws and errors. For example, PSE noticed in PEG's working papers that PEG incorrectly subtracted certain Hydro One costs when attempting to align the cost definitions between Hydro One and the U.S. sample. In PEG's response in EB-2019-0082, L1, Tab 1, Schedule 21 (a) and (b), PEG acknowledged and corrected this error, although it has not revised the PEG Report in this regard. The correction changed Hydro One's 2020 to 2022 benchmark score to +6.8% from +9.0%.

Beyond PEG's cost definition error, this section discusses the two major flaws in PEG's research that will have a large impact on PEG's benchmarking results. PEG put forth several concerns in its report on PSE's research that are inconsequential. In contrast, we are focusing here on only the two concerns that will have a major impact on the results (although in our view there are other, more minor errors in PEG's approach). If these two major errors are fixed, the PEG model would show that Hydro One is a strong cost performer, indicating a stretch factor of 0.0% consistent with PSE's analysis. These two major flaws are:

1. PEG's model contains a clear and obvious bias against the recent years (and the forecasted years) for all utilities in the sample. This unfair bias has a major impact on PEG's evaluation of Hydro One's custom IR period of 2020 to 2022. When corrected, Hydro One's performance improves significantly.

2. PEG instituted a different modeling procedure that drastically changed its reported results from only six months prior. This modeling procedure is not transparent and is open to subjective decisions by the researcher. The PEG procedure also possibly contains errors and, even if it was instituted properly, does not offer any statistical improvement over PSE's method. There was no need for this change.

2.1 PEG's Results are Biased Against the Recent and Forecasted Years for All Utilities in the Sample

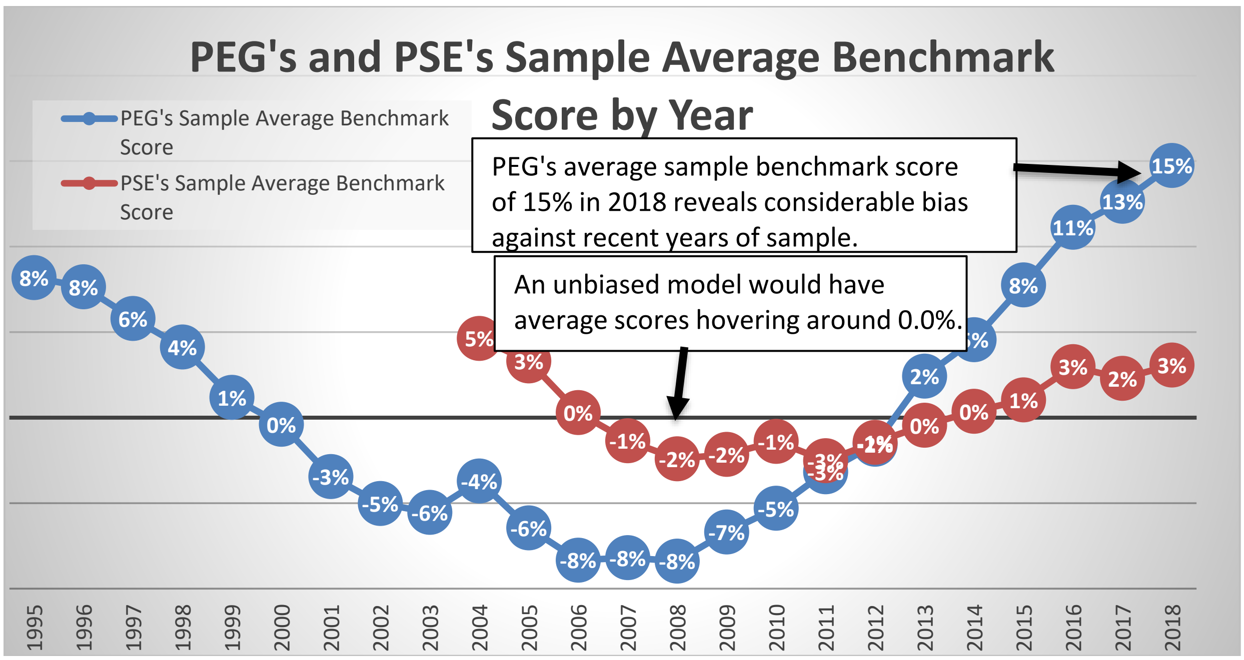
The first major flaw in the September 2019 PEG Report is this: PEG's model has a serious bias against the more recent years in the sample. This bias is present across the entire sample. This is why the PEG results erroneously show a precipitous drop in Hydro One's cost performance from 2004 to 2016, despite PEG's finding that Hydro One's productivity outpaced the industry during that same period.[18]

[18] In Table 3 and 4 of the PEG Report, PEG shows that Hydro One's TFP trend is 0.3% higher than the industry during the 2005 – 2016 period.

This counterintuitive result is due to PEG's model being biased during the more recent periods of the sample period. The bias is clearly present for the entire sample. PSE inserted the 2017 and 2018 observations into PEG's dataset and calculated the benchmarking scores for each observation in each year. The figure below provides the average benchmark score for the sample for both the PSE and PEG samples in each year.[19] The blue line shows the bias in each year for the PEG sample. The red line shows the bias in each year for the PSE sample.

We would expect a model without a systematic bias to have sample average scores that hover around 0%. We would expect 0% to be the average benchmark score for the sample because this would indicate an average-performing utility is at their benchmark (or expected) total costs. The objective of performance benchmarking is to provide a comparative analysis showing how a target utility's costs compare to a hypothetical average utility sharing the same characteristics as the target utility. PEG's model is not producing those results but, instead, is calculating the benchmark scores of the entire sample to be 15% higher in 2018.[20]

Figure 1 PEG's and PSE's Sample Average Benchmark Score by Year



There is a clear trend in PEG's average benchmark score. The PEG benchmark scores exhibit a curved (or quadratic) trend.[21] The PSE results hover around the expected level of 0%. As is clear from the "U" shape in the blue line above, the bias in PEG's model will continue to grow through

[19] Recall that a benchmark score is the percent difference between the utility's actual total costs and its benchmark total costs. We would expect an average-performing utility to have a benchmark score of 0.0%, indicating its total costs are the same as its benchmark costs.

[20] Hydro One asked PEG to provide the benchmark scores for the sample in 2014, 2015, and 2016. PEG provided these in EB-2019-0082, L1, Tab 1, Schedule 23 (a) in an attachment. In 2014, PEG reports a sample average benchmark score of 4.7%, in 2015 it increases to 7.8%, and in 2016 it increases again to 11.2%.

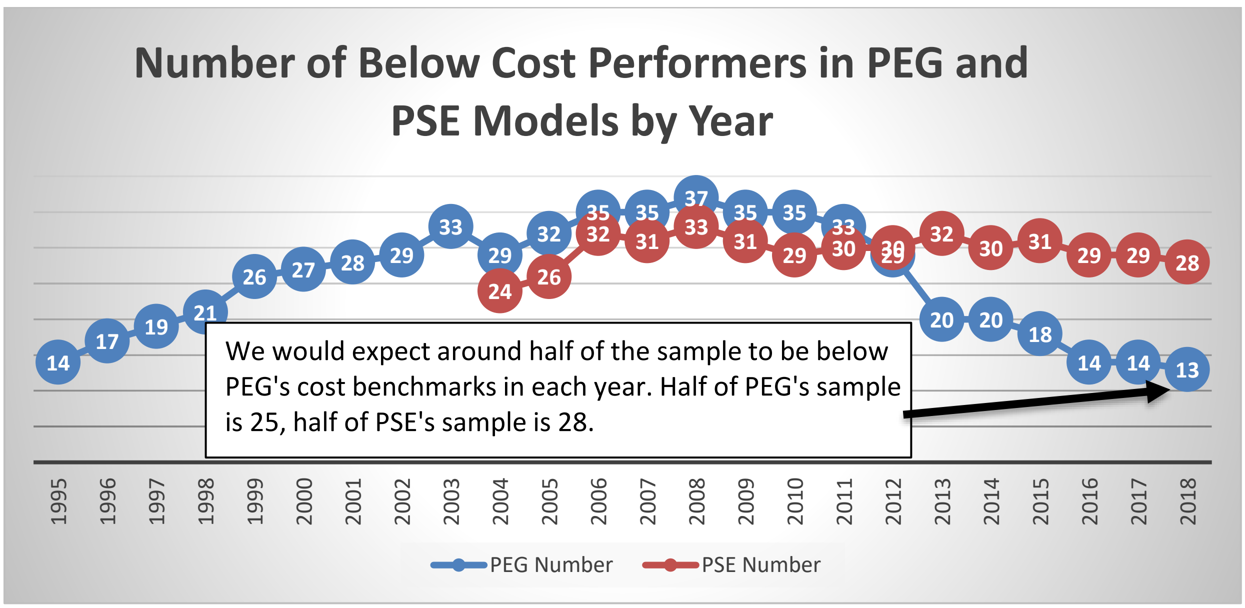
[21] The benchmark scores are the residuals of the econometric model. When a clear pattern is present in the residuals it indicates the model is not specified properly.

the year 2022. This bias is significantly and unfairly harming Hydro One's benchmark score in the recent and forecasted years of the sample.

The following graph further illustrates the first major error in PEG's approach. PEG's model includes 50 utilities in 2018. If PEG's results are normally distributed, one would expect around half of the utilities (i.e., 25) to be in the "below cost category" in each year. Instead of a number close to 25, PEG's model has 37 utilities deemed to be below cost in 2008, and only 13 deemed to be below cost in 2018, around 25% of the sample. In PEG's modeling approach, as the year approaches 2018, it gets harder and harder for any utility in the sample to be deemed a low-cost utility. Because of this distortion, even though PEG's results indicate a 0.3% stretch factor, Hydro One is right at the border of a <u>top quartile utility</u> in 2018 (13th out of 50) after PEG corrected for its cost definition error in L1, Tab 1, Schedule 21 (b).

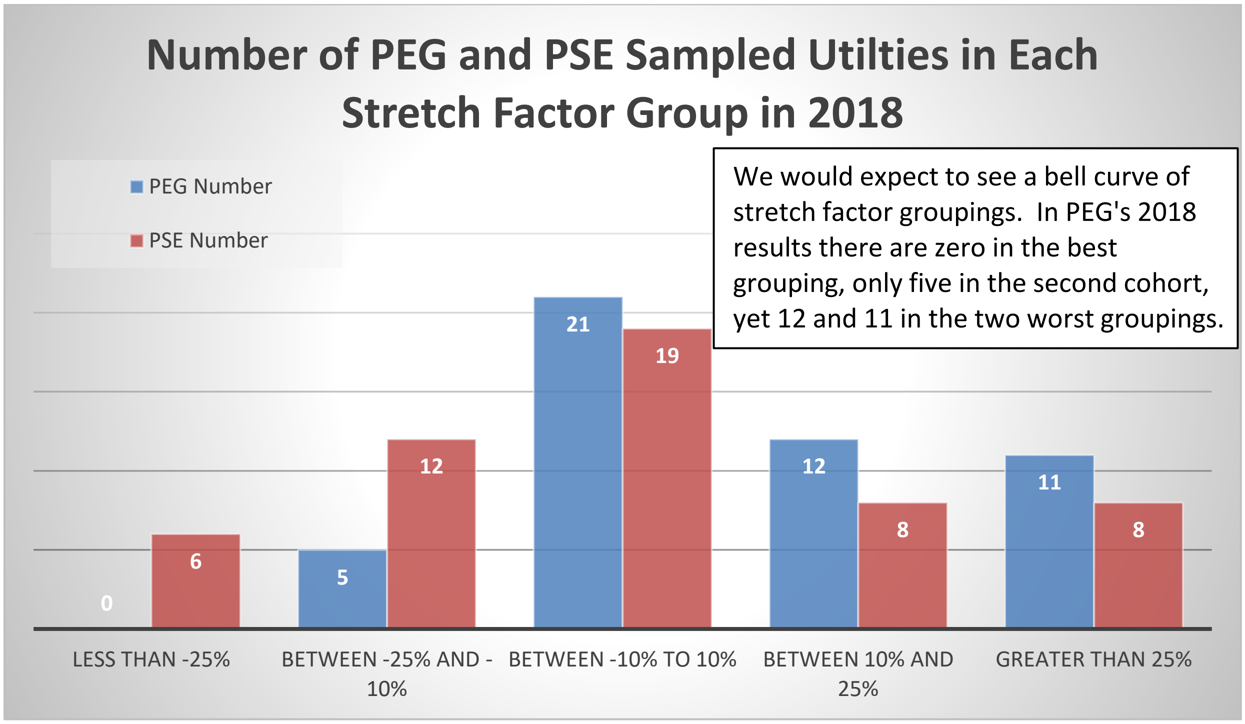
PSE's model, by comparison, shows a much more predictable and consistent number of utilities above and below the benchmarks. PSE's sample includes 56 U.S. utilities. We would expect around 28 utilities to be in the "below cost category" in any given year. PSE's results (red line) show this consistency, hovering around 28 in each year.

Figure 2 Number of "Below Cost" Performers in PEG and PSE Models by Year



Another way to illustrate the bias is to see the spread of the benchmark scores and what the benchmark scores in 2018 would indicate as far as a stretch factor. We see in the figure below that PEG's model would produce zero utilities in the 0.0% stretch factor cohort for 2018, but would produce eleven in the worst cohort. A comparative benchmarking analysis should be close to symmetric and have a "bell curve" shape in the number of utilities deemed high cost, average, and low cost. PEG's results do not exhibit this bell curve in the recent years of the sample. In 2018, PEG only has 5 utilities that would land in the best two cohorts (0 in the 0.0% cohort and 5 in the 0.15% cohort) but has 23 utilities in the worst two cohorts (12 in the 0.45% cohort and 11 in the 0.6% cohort). In comparison, PSE's results do exhibit a bell curve, with 18 utilities in the best two cohorts and 16 in the worst two.

Figure 3 Number of PEG and PSE Sampled Utilities in Each Stretch Factor Group in 2018



The bias inherent in PEG's benchmarking results can be further illustrated by looking at the year 2019 and PEG's own research results. In Table 4 of the PEG Report, PEG shows the productivity trends of Hydro One from 2004 to 2022. PEG calculates a TFP increase for the company in 2019 of 1.00%. However, PEG's total cost benchmarking score for the company increases by 1.0% in that same year. A reasonable benchmarking model would not have Hydro One getting a "worse" score in a year when the company's productivity outpaced the industry's by well over 1%.

How could a utility ever improve in PEG's benchmarking model in future years given that its score gets worse by 1% even when productivity exceeds 1%? A utility would need to have a sustained productivity trend of over 2% just to not get worse in PEG's model in the forecasted years. This is not a reasonable outcome.

2.1.1 Simple Fix of Adding One Variable

PEG's model is clearly biased against the later years in the sample. There is a simple fix to PEG's flaw that only requires the addition of one variable, and no other changes. If a quadratic trend variable is inserted into PEG's model to capture the curvature of the cost trends, the model will be far better at accurately predicting cost levels, and the variable will substantially reduce the bias against the recent and forecasted years.

A quadratic trend variable is a reasonable variable to insert when real cost trends exhibit the "U" shape that is clearly observed in PEG's chosen sample period that begins in 1995 (see the "U" shape in Figure 1 above). Dr. Lowry of PEG has mentioned that a quadratic trend variable is a reasonable variable to consider in the recent Toronto Hydro application (EB-2018-0165). Here is an excerpt from Dr, Lowry's testimony in the Toronto Hydro hearing on July 15, 2019 (p. 43 of the transcript).

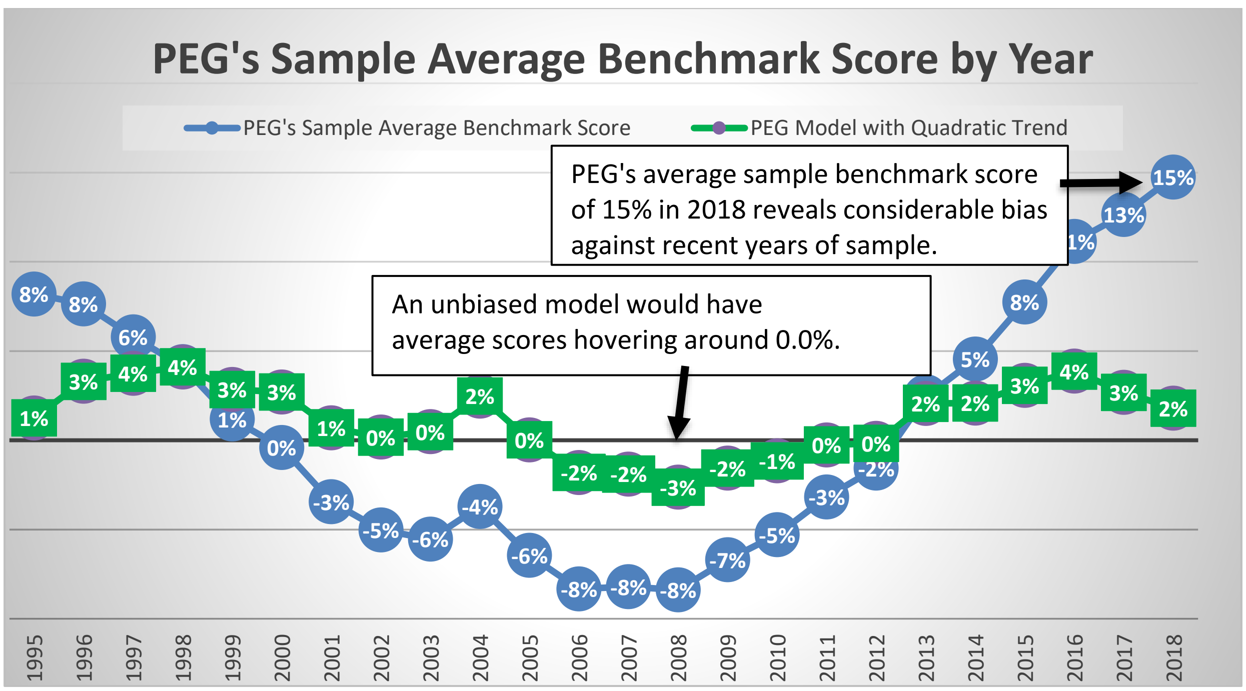
Figure 4 PEG Testimony

MR. SHEPHERD: A quadratic for the trend variable would change the trend variable from a straight line to curved line, correct?

DR. LOWRY: Yes. I mean, why not have that? I mean, I would be more interested in that to some degree than in some of the others. Particularly when you are forecasting outside of the sample period, it might be interesting to have a curvature on that.

When a quadratic trend variable is inserted in PEG's model, and with no other changes made, PEG's bias in each year hovers around the expected 0% value. In 2018, the bias is only 2%. The following graph displays the bias in PEG's reported model (blue line) and PEG's model with the only change made being the insertion of a quadratic trend variable (green line).

Figure 5 PEG's Sample Average Benchmark Score by Year with Quadratic Trend



By including the quadratic trend variable into the PEG analysis and leaving all other methods the same, we estimate that PEG's Hydro One benchmarking scores for the 2020-2022 period will be lowered from PEG's reported +9.0% score by 25.1%: this one variable addition, with no other changes to PEG's methodology, results in a PEG benchmark score in 2020 to 2022 of -16.1%.[22,23]

2.2 PEG's Different Modeling Procedure is Not the Proper One to Use

The second major flaw in PEG's analysis in its September 2019 Report was to change and complicate the modeling procedure, for no convincing reason. PEG changed its modeling procedure in the present case, relative to what it did in the HOSSM and the Hydro One Distribution cases.[24] Compared to the previous cases, PEG has now coded into their customized econometric code an adjustment for autocorrelation, named a Prais-Winston adjustment. This adjustment is in addition to the prior adjustment PEG coded to address heteroscedasticity. The code written by PEG and its underlying assumptions needlessly influence the coefficient values of the model.[25] This change has had a considerable impact on PEG's benchmark results for Hydro One, compared to the results that PEG reported approximately six months prior.

In the HOSSM case, PEG corrected certain errors discovered by PSE in PEG's response to interrogatory PEG-HOSSM-6i. In an attachment labeled "Attachment PEG-HOSSM-6i(b)" to that response, PEG displayed a table showing that Hydro One's 2014-2016 average total cost score was -22.87%, and that its 2019-2022 average total cost score was -12.35%. Below is the table produced by PEG in the HOSSM case.

We note that PEG now claims the HOSSM modeling procedure was not valid in EB-2019-0082, L1, Tab 1, Schedule 24 (b), although PEG stood behind its model and its work during the HOSSM proceeding. In the Hydro One Distribution proceeding in EB-2017-0049, L1, Tab 8, Schedule 53, PEG stated they did not conduct an autocorrelation correction.

[22] We note that PEG was requested to add the quadratic trend variable to their model and provide the results, but refused this request in their response to EB-2019-0082, L1, Tab 1, Schedule 6, part h.

[23] If we consider the effects of both: (1) The quadratic term mentioned in this section, and (2) the correction Hydro One's cost definition mentioned in the beginning of this Section 2 , Hydro One's 2020 to 2022 benchmark score would be -18.3%. PEG's reported score for this period was 9.0%; lower the score by 25.1% due to the quadratic term; lower by 2.2% for the cost definition; to end at -18.3%.

[24] PEG stated in EB-2017-0049, Exhibit L1, Tab 8, Schedule HONI-53 that it did <u>not</u> conduct an autocorrelation adjustment in its research of Hydro One Distribution. This statement is contradicted by PEG's statement in this case in L1, Tab 1, Schedule 24, part c and d where PEG claims it <u>did</u> conduct an autocorrelation adjustment in its Hydro One Distribution research.

[25] PEG's complex adjustments require customized coding by PEG and cannot be replicated by any off-the-shelf software that we are aware of without requiring the researcher to code in the procedures.

Figure 6 Hydro One's Total Transmission Cost Performance Reported by PEG in HOSSM

Attachment PEG-HOSSM-6i(b)

Hydro One's Total Transmission Cost Performance Using PEG's Model

[Actual - Predicted Cost (%)][1]

Year	Cost Benchmark Score
2004	-41.20%
2005	-44.20%
2006	-43.30%
2007	-38.50%
2008	-41.00%
2009	-34.70%
2010	-32.40%
2011	-31.80%
2012	-27.90%
2013	-25.30%
2014	-25.00%
2015	-21.60%
2016	-22.00%
2017	-20.50%
2018	-18.70%
2019	-16.40%
2020	-13.70%
2021	-11.00%
2022	-8.30%
Average 2004-2016	**-32.99%**
Average 2014-2016	**-22.87%**
Average 2019-2022	**-12.35%**

[1] Formula for benchmark comparison is $\ln(Cost^{HOSSM}/Cost^{Bench})$.

However, in the present case, in Table 5 on p. 38 of the PEG Report, we see a substantial change in PEG's benchmarking results for Hydro One Networks. PEG's results have now changed to -2.1% for the 2014-2016 period, and +9.0% for the 2020-2022 period. After examining PEG's working papers, we have discovered the primary cause of this change is PEG coding in and implementing a different modeling procedure from what PEG used in the HOSSM proceeding.

PEG agrees the modeling change is a large contributor to the modified results.[26] PEG acknowledged that the results changed more than one might expect. PEG appears to justify this change by characterizing Hydro One's business conditions as "atypical".

Beyond the outputs (line length and maximum peak demand), which PEG says had a minimal impact from the change in modeling procedures, the only large anomaly in Hydro One's business conditions is the percent transmission variable used by PEG. PEG claims this variable is almost five times the sample average. However, PEG described this variable incorrectly when they stated:

[26] See EB-2019-0082, L1, Tab 1, Schedule 1 (a).

"Hydro One is the only company in the sample that only performs transmission service." This statement is clearly not true. PEG inserted the incorrect value of "100%" in its dataset for this variable. This incorrect value inserted by PEG is the reason that PEG calculated the company's percent transmission variable as being nearly five the sample average.

The company is much less of an outlier if PEG had taken the correct approach that PSE undertook, and accounted for the fact Hydro One has both transmission and distribution operations. If PEG had calculated the proper percent transmission variable, Hydro One's benchmark score would have increased by around 6%.

2.2.1 Why PEG's Changed Modeling Approach is Flawed

We have investigated the changed PEG modeling procedure and have concerns on PEG's changed approach, beyond the fact that it is not consistent with their own prior work.[27] This method significantly and needlessly altered the benchmarking results relative to what PEG produced in HOSSM and relative to what a more modern, reproducible, less subjective, and transparent econometric approach would indicate. PSE's model procedure is the more modern approach. The key advantage of the PSE method is that it directly corrects the problems associated with heteroskedasticity and autocorrelation, without manipulating the model coefficients that formulate the benchmark scores.[28]

Two common problems arise in econometric modeling using real world data: heteroscedasticity and autocorrelation. When present, heteroskedasticity and autocorrelation can decrease or increase the regression standard errors associated with each coefficient value. It is important to note that neither of these problems causes the coefficient values to be biased. In other words, the researcher does not need to worry about correcting the coefficient values for any problems or biases: the coefficients are not misleading, and they cannot be improved upon. And it is these coefficient values that are used to calculate the benchmarks.

The modern view that is becoming more standard is that these coefficient values should be left alone and not manipulated. PEG's approach changes the coefficient values based on the researcher's underlying assumptions of what is driving the heteroskedasticity and autocorrelation. PSE's modern approach does not modify the coefficient values used to calculate the benchmark scores and requires no assumptions by the researcher.

There are several correction methods designed to increase the precision estimate of each standard error caused by autocorrelation and heteroscedasticity. What has become a standard approach in econometrics is to choose a method designed to only correct the standard errors while leaving

[27] Dr. Kyle Stiegert, an economics professor at the University of Wisconsin-Madison in the Agricultural and Applied Economics Department assisted with this investigation and this discussion on why the PSE approach is the preferred one. Dr. Stiegert has taught graduate level econometric courses and authored dozens of journal articles applying econometrics to real-world contexts.

[28] In statistics, heteroskedasticity happens when the standard errors of a variable are non-constant. Autocorrelation is a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals.

untouched the ordinary least squares (OLS) coefficient values (see Wooldridge, sections 8.2 and 12.5).[29] This approach is commonly referred to as robust standard errors. Wooldridge (2012) states "In large sample sizes, we can make the case for always reporting *only* the heteroscedasticity-robust standard errors in cross-sectional applications, and this practice is being followed more and more in applied work."[30] A basic and commonly applied robust method for confronting both heteroscedasticity and autocorrelation was developed by Newey and West (commonly called Newey-West standard errors). However, the Newey-West correction cannot be used for unbalanced panel datasets. The Driscoll-Kraay method that PSE uses produces robust standard errors that are corrected for both heteroskedasticity and autocorrelation. The Driscoll-Kray method was developed for use in unbalanced panel datasets like the data used in our analysis.

Before the advances in robust standard errors (described above), econometricians would attempt to adjust the standard errors using weighted least squares (WLS) methods that could also substantially alter the coefficient values. WLS requires the researcher to assume what weightings should be used to make the adjustments. PEG employs a WLS method to correct for heteroscedasticity (called panel corrected least squares), and then uses a second correction (called Prais-Winston) to purge autocorrelation.[31] Given the considerable change in PEG's results now compared with its HOSSM results of about six months ago, it is clear that the WLS methods and PEG's underlying assumptions can substantially alter the forecasts, compared to the base case established by the standard OLS coefficients. These alterations are made by PEG, despite there being no ability to improve the OLS predictions; the alterations can only harm the predictions (this could occur if PEG's underlying assumptions are not accurate, the adjustments should not be made on an unbalanced panel dataset, or if PEG codes the complex adjustments incorrectly).

The method that PSE employs (Driscoll-Kraay or DK) does not in any way allow the researcher to manipulate the coefficient values that drive the forecast results. As a modern advancement for confronting the problems of heteroscedasticity and autocorrelation, the coefficient values from our analysis are the OLS values and they cannot be improved upon.

The only information that moves forward from the regression step to the benchmark calculation step are the coefficient values. As stated earlier, PEG's WLS correction procedures changes the coefficient values from OLS, which in turn changes the benchmarks. WLS requires assumptions on the underlying causes of heteroskedasticity and autocorrelation; if those assumptions are incorrect, PEG's coefficients and the accuracy of the benchmarks will be adversely impacted. A researcher using the PEG approach can select from various options in the WLS framework and possibly choose the preferred forecast for their client. PSE's correction method removes this ethical dilemma by maintaining the OLS estimates.

[29] Wooldridge, Jeffrey M. 2012. *Introductory Econometrics: A Modern Approach, 5th edition*. South-Western Cengage Learning. United States.

[30] *Id.* p. 273 (italics added for emphasis).

[31] We note that PEG only made the first WLS adjustment in their HOSSM research and this first adjustment had a minor impact on the model coefficients and resultant benchmarks relative to OLS and PSE's method. It is the Prais-Winston adjustment that PEG coded and instituted now that has significantly impacted the model coefficients and the resultant benchmarks.

Given the complexity and customization of PEG's econometric coding, we are unable to verify that PEG's new modeling procedure has been coded properly or that performing these two WLS adjustments on an unbalanced panel dataset is a valid approach to begin with. A more transparent and reproducible method would be to use commercially available econometric software packages where the procedures are coded by the vendor and verified by the public use of those procedures.[32] However, PEG has not provided any commercially available software packages that can replicate PEG's results that do not require the consultant to customize the code.

We are forced to assume without verification that PEG conducted all its complex coding properly, these procedures are valid, and made reasonable assumptions on the underlying sources of heteroskedasticity and autocorrelation when making these needless adjustments. This assumption is made more difficult when PEG's results change so drastically from six months ago due primarily from them coding and implementing a different and complex modeling procedure.

In EB-2019-0082, L1, Tab 1, Schedule 24 (b), PEG confuses the efficiency of the standard errors and coefficient estimates. Recall that the issue with autocorrelation and heteroskedasticity is with the standard errors, not with the coefficient estimates that produce the benchmarks. PEG's modeling approach does <u>not</u> produce better coefficient estimates than PSE's DK method that uses the OLS estimates for the coefficient estimates. Further, PEG's approach does <u>not</u> produce more efficient standard errors than PSE's DK method that adjusts the standard errors for heteroskedasticity and autocorrelation. The PEG coefficient estimates are of equal quality with PSE's only <u>if</u> all of the following are true : (1) PEG's underlying assumptions are correct, (2) the two procedures are valid procedures to undertake to an unbalanced panel dataset, and (3) the procedures were coded properly. We cannot verify these three conditions.

In summary, we believe that PEG's changed modeling approach:

(1) Introduces a possibility for error, given the complex coding necessary to undertake its new modeling procedure,

(2) Has not proven to be a valid procedure on an unbalanced panel dataset,

(3) Is not necessary since the coefficients from the standard OLS run cannot be improved upon,

(4) Is open to subjective judgement by the researcher, and

(5) Is not easily reproduced and verified by non-experts using standard econometric software packages.[33]

[32] PSE's results can easily be reproduced by several commercially available econometric software packages with no customized coding required. In fact, in our working papers we provided results from two such vendors (EViews and STATA).

[33] While PEG's modeling procedure requires extensive customized code to be written with little ability to identify errors, PSE's benchmarks can be replicated by most off-the-shelf econometric software packages.

By improving PEG's modeling procedure to use the DK procedure which uses the OLS coefficient estimates and making no other changes, we estimate that PEG's Hydro One benchmarking scores for the 2020-2022 would be lowered by 29.5% and become -20.5%.[34]

2.3 Summary of Estimated PEG Results When Corrections Are Implemented

PSE was able to re-run PEG's reported results using its code (although as noted above, we cannot verify some of PEG's decisions, such as whether PEG has coded adjustments properly or if these adjustments are valid ones to use). We then fixed PEG's errors one-by-one to see the impact of each change on the PEG results. The table below provides PSE's estimates of those impacts.

Table 3 Impact on PEG's Hydro One 2020 to 2022 Score When Corrections are Made

Methodology	Estimated Impact on Hydro One Average 2020-2022 Score with Corrections in PEG Method and No Other Changes Made
Correction 1: Reduce clear bias in PEG's model against the recent and forecasted years of sample	-25.1%
Correction 2: Use OLS coefficients that are not open to manipulation, don't require assumptions, and are far more transparent (or revert to PEG's HOSSM and Hydro One Distribution modeling approach)	-29.5%

We note that these two major corrections; 1) using a quadratic trend variable, and 2) using OLS coefficients or, at the very least, PEG's HOSSM approach, each have a large impact on Hydro One's score. If either one of these corrections is implemented, PEG's results would indicate a 0.15% stretch factor. If both are implemented, PEG's results would indicate a 0.0% stretch factor, and would be consistent with PSE's analysis. Both are corrections which should be made.

3 Reply to PEG Concerns

Starting on p. 19 of the PEG Report, PEG raises some concerns it has with our productivity and benchmarking studies. In respect of each stated concern, we either disagree with PEG or note that it is inconsequential to the study results. We provide our replies below to each point raised by PEG.

3.1 The Productivity Study

PEG states three concerns on PSE's productivity study.

[34] The original score was 9.0%; if improved by 29.5%; the score becomes -20.5%. If we account for PEG's corrections to its error in subtracting Hydro One's costs (2.2%), Hydro One's benchmark score becomes -22.7%.

1. Sample Period
2. Structural Change
3. Capital Cost Specification

3.1.1 Sample Period

PSE and PEG's results are quite similar when examined over the same sample period. For example, PEG's 2005 to 2016 industry TFP trend is -1.47%, and PSE's TFP trend is -1.45%. Likewise, PEG's 2011-2016 TFP trend is -2.33% and PSE's TFP trend is -2.39%. This shows that the main difference in the results is the chosen sample period of the study.

In EB-2019-0082, L1, Tab 1, Schedule 6, PEG was asked to update its analysis to 2018, but PEG refused this request. Given PEG's concerns regarding the length of the sample period,[35] PSE has now added two years to its sample period, and our sample now includes 14 years containing the most recently available information. The 2005 to 2018 industry TFP trend is -1.61%.

Examining the 2017 and 2018 TFP results illustrates why the PSE sample period is more appropriate in setting productivity expectations for the upcoming years of the Custom IR plan. Both the PSE and PEG productivity results reveal there is a consistent and pronounced slowdown in productivity in recent years. In the last 10 years of PEG's sample (2007 to 2016), every year had productivity below PEG's recommended TFP finding of -0.25%. The years 2017 and 2018 followed this clear trend by showing productivity declines of -1.5% and -3.4%, respectively. This makes 12 consecutive years where PEG's recommended TFP trend would have overestimated the realized trend. PEG has shown no evidence to suggest these negative productivity trends will abate in the next few years.

The TFP sample period should consist of at least the most recent ten-year period. However, going further back in time is not necessarily desirable. Data considerations, technology changes, industry expectations, output growth, and structural changes should all be considered. Given the large structural change in the industry attributable to the move to Independent System Operators (ISOs) that occurred in the late 1990's and early 2000's, the increase in distributed energy resources (DERs), the slowdown in output growth, and the aging infrastructure issue within the electric industry, beginning the sample period in 2004 will be far more reflective of the expected productivity experience in upcoming years than PEG's sample that begins in 1996.

PEG's TFP trend of -0.25% is heavily influenced by the strongly positive TFP trends of the 1990s. These trends are not applicable to today for the following reasons.

1. **Output growth** is far different now then back in the 1990s, especially for Hydro One. Hydro One is projecting near zero output growth during the Custom IR period. The growth of DERs throughout the grid have also lowered output growth and slowed TFP trends. PEG finds on Table 3 of its report that output growth for the industry increased by over 1% per year during the 1990s but the industry growth has now slowed considerably. Hydro One's output growth is projected to be 0.0% for the Custom IR plan.

[35] See p. 19-22 of the PEG Report.

2. The **structural change** towards ISOs/RTOs that the industry underwent in the late 1990s and early 2000s is a structural change that should not be included in the TFP sample period because there is no anticipated similar change in the industry in the upcoming years of 2021 and 2022. A sample period occurring after this profound structural change is preferred to formulate an appropriate expectation of the 2021 and 2022 TFP trends.

3. The **aging infrastructure** issue was far less of a challenge back in the 1990s. Due to the post World War II baby boom and the increased electrification of society, electric utilities invested heavily in new infrastructure during the 1960s and 1970s. These investments were funded by the output growth of the industry. However, this output growth is no longer present today, and these assets are now 40 to 60 years old. This situation was far less of an issue back in the 1990s, when the "baby boom" assets were only 20 to 40 years old. PEG agrees with this as it states in its response to EB-2019-0082, L1, Tab 1, Schedule 12, part b that, "PEG does not believe that the challenge of aging industry infrastructure is likely to be rectified by 2021."

4. As PEG mentions in its report, there has also been an increased focus on transmission grid **reliability** since the 1990s. Added to that are new concerns such as **cybersecurity**. These concerns have increased since the 1990s and are not likely to abate in the near-term.

PSE's sample period now consists of 14 years and is the best available measure to base the productivity factor for Hydro One's Custom IR period of 2020 to 2022. The two additional years confirm the declining productivity trend exhibited in recent years. The 2005 to 2018 results show an average annual TFP decline of -1.61%. The 14-year time length is comparable to other studies used to formulate an X-factor. In 4GIR, the sample period used for the electric distributors was from 2003 to 2012 (i.e., a ten-year period). In the HOSSM application, PSE mentioned in an interrogatory response that one of the primary studies we reviewed was a recent electric transmission study from the Australian Energy Regulator (AER).[36] The AER study's time period dated from 2007 to 2016, a ten-year period. Similar to both the PSE and PEG results for US productivity, the AER found declining TFP during this time period.

In PEG's report in the amalgamation application between Enbridge Gas and Union Gas (EB-2017-0306/EB-2017-0307), PEG presented productivity evidence in that case and filed a report (Exhibit M1). On p. 42 and 43 of that report PEG discusses the appropriate sample period for a productivity study. The criteria stated by PEG are:

1. Include the latest year for which requisite data is available
2. Sample period should reflect the long-run productivity trend, so it is desirable for the sample period to be at least ten years in length.
3. A long sample period, however, may not be reflective of the latest technology trend.
4. The start date for the period should be several years after the capital benchmark year.

PSE's sample period of 2005 to 2018 accomplishes all four of these criteria.

[36] EB-2018-0218, Exhibit I, Tab 1, Schedule 63, p. 2 of 4.

1. The PSE sample period includes the latest available data for the years 2017 and 2018. PEG's sample ends in 2016.

2. The PSE time period comes after the majority of the ISO/RTO structural changes of the late 1990's and early 2000's occurred. PEG's sample includes this large structural change during its sample period.

3. The PSE sample period, while being a robust 14 years long, does not dilute the clear recent changes in TFP trends possibly due to aging infrastructure, slowing output growth, and increased reliability and security demands on transmission systems. PEG's sample period does dilute the clear TFP trends of recent years by inserting observations that were during a far different period of faster output growth, newer assets, and lower reliability and security concerns.

4. PSE's sample period begins in 2004 which is 15 years after the capital benchmark year of 1989. This is a sufficient gap to ensure the capital costs are being properly accounted for.

3.1.2 Structural Change

PEG's concern over the ISO/RTO structural change impacting PSE's research is unwarranted, and instead should be directed at PEG's own choice of sample period. The move to the ISO/RTO transmission industry structure occurred during the late 1990s and into the early 2000s. As PEG states on p. 72 of the PEG Report, "Several ISOs were formed between 1996 and 2000." This industry structural change is within PEG's 1996 to 2016 sample period and could have a strong influence on PEG's results.

In the HOSSM case, Hydro One asked PEG how many utilities in its sample transitioned to ISOs/RTOs during their longer sample period. PEG's response in EB-2018-0218, L1, Tab 1, Schedule 8, showed that during PEG's sample period 39 utilities joined an ISO/RTO. This is well over half of PEG's TFP sample. In contrast, PSE's sample period only included 6 utilities that joined an ISO/RTO. We are of the opinion that a sample period that begins after this structural change is the more appropriate time period to utilize for both the TFP and benchmarking studies when formulating forecasts for a period that will not contain this structural change.

3.1.3 Capital Cost Specification

PEG's concern here is inconsequential. PEG itself demonstrated in HOSSM that changing the capital benchmark year from 1964 to 1989 would have a small impact on the benchmarking results.[37] PEG states in Exhibit L1, Tab 1, Schedule 5, part b that PEG has no reason to believe the impact would be larger now.

We also note the similarity in the TFP trends of both PSE and PEG when the same sample periods are examined. The reason that the issue is trivial and should not have been raised by PEG is because the capital additions occurring from 1965 to 1988 are substantially depreciated by the sample years. Further, any differences from beginning the capital series in 1964 or 1989 are

[37] EB-2018-0218, PEG-HOSSM-6j.

reflected through the entire TFP sample period, so the differences will have a minimal impact on the estimated TFP trend.

PEG's filing in other cases show that they agree with us on this point. In a report filed on behalf of Public Service Company of Colorado's gas utility, dated May 31, 2017 and titled "Statistical Research for Public Service Company of Colorado's Multiyear Rate Plan," PEG's productivity and benchmarking research used a capital benchmark year of 1984 and had a start date in their sample of 1998, a 14-year difference. On p. 44 PEG states: "Any inaccuracy in these assumptions is mitigated by the fact that plant additions from years before 1984 are substantially depreciated by the later years of the sample period."

In this case, PSE uses a capital benchmark year of 1989 and begins the sample in 2004, a 15-year difference. If the statement by PEG of their own work in Colorado is accurate (and we believe it is), then PSE's work in this case, using a capital benchmark year of 1989, is similarly an inconsequential concern.

There also exists a likelihood of increased errors when using the older data going back to 1964. PEG refuses to provide the source data so others can readily review PEG's dataset, despite the data not being electronically available. PEG admits it was gathered "decades ago," and that source book titles cannot be named in each year.[38] Unlike all of PSE's capital data, this older data is not electronically available and would require an immense effort on PSE's end to track down and gather. It must be manually entered, with human error likely to occur. In fact, in HOSSM we saw this first-hand, when PSE identified inconsistencies in this older capital data between PEG's TFP and benchmarking studies. This caused a significant change in PEG's total cost benchmarking results for Hydro One and pushed PEG's benchmark scores for the Custom IR period to -11% for Hydro One, which would imply a 0.15% stretch factor.

This high likelihood of and history of errors and large manual process required when using this older data, the fact it was gathered decades ago, and the refusal to not allow a third party to verify the data, far exceeds any possible slight increase in accuracy it may offer in our view.

3.2 The Benchmarking Study

PSE disagrees with, or finds inconsequential, each of PEG's concerns regarding our benchmarking study. The big picture is that variables between the PSE and PEG models are almost the same, except for one difference when PEG leaves out one obvious variable (# of transmission substations).[39] The biggest differences in methodologies that impact the results are the two PEG flaws that we discussed in Section 2.

3.2.1 Sample Period

Please see Section 3.1.1 for an overview of our opinions regarding the sample period and our TFP research. That discussion is applicable to the benchmarking sample period as well. See also Section

[38] See PEG's response to EB-2019-0082, L1, Tab 1, Schedule 5 part e and f.

[39] This omission, however, did not have a consequential impact on Hydro One's results so we did not mention it in our suggested model corrections found in Section 2.

2 where we demonstrated that PEG's modeling approach produces highly biased results for the entire sample, including Hydro One, in the most recent years of the sample period. PSE's time period enables our model to better reflect the current parameter values and transmission cost drivers that best capture the impact of variables onto costs in recent and forecasted years.

Fundamentally and as discussed above, PSE's shorter sample period enables our model to have much less bias than PEG's approach. The benchmarking results are mainly used to examine Hydro One's recent and projected total cost performance, and therefore including observations from the 1990s is not helpful. Technology advances, infrastructure age, slower output growth, regulatory requirements, ISO/RTO transition, and reliability expectations have evolved throughout the years, and thus a more contemporary sample is more reflective of the current conditions and reality within the industry.

3.2.2 The Trend Variable Parameter

PSE's trend variable is far more reflective of the current industry trend than PEG's. As adding the 2017 and 2018 years demonstrated, PSE's trend variable is a better reflection of upcoming cost trends. In Section 2 we showed that the PSE model contains far less bias and is a better predictor of total cost levels. Given the similarity in included variables, one of the biggest reasons for our model being better is that PSE's trend variable appropriately captures the trends in costs, whereas PEG's does not.

PEG's benchmarking model is assuming a positive productivity trend and costs to increase below inflation for future years. This is inconsistent with PEG's own research. PEG's research shows productivity trends below -2% in recent years. PEG's assumption would have been inaccurate for the last 12 years, including 2017 and 2018 years.

PEG's trend variable is creating a severe bias in the results for the recent and forecasted time periods (see Figure 1 in Section 2). While an unbiased model would show an average benchmark score close to 0.0%, indicating that an average-performing utility would be at its benchmarks, PEG's bias in recent time periods is large and increasing over time. This built-in bias against recent and future years for all utilities in the sample is the reason that Hydro One's benchmarking scores are declining over time despite Hydro One's measured TFP being higher than the industry in both PSE and PEG's calculations. PSE has no such systematic bias in the recent and forecasted time periods due to the more appropriate time period used by PSE.

3.2.3 Capital Data Starting in 1964

We view this as an inconsequential item that should not be a concern of the Board. Please see our prior comments in section 3.1.3 for our full explanation of the inconsequential impact of using this data, and why using the 1989 data is preferable, due to the data being electronically available with a far lower likelihood of manual entry errors.

3.2.4 Hydro One's Capital Series Starts in 2002

We agree with PEG that this is beyond the control of both PSE and PEG. Both models start the capital series for Hydro One in 2002 due to the data limitations. This will reduce the accuracy of

Hydro One's TFP and total cost benchmarking scores in the earlier years of the sample period however, that concern should cease to apply by the later years of the sample period.

3.2.5 Construction Standards Index Variable

PSE took the proper approach that is consistent with how the US sample is calculated. This approach used the service territory of each company, including Hydro One, to formulate the variable. If it were possible to switch to one based on calculating the variable based on the location of transmission lines for each company, the benchmarking results would likely improve for Hydro One. Despite PEG citing this as a concern, PEG refused to revise their results with a variable value that addressed these concerns.[40] PSE estimates that modifying the dataset to reflect PEG's concern would improve Hydro One's benchmarking score by about 3.5%. However, we do not believe this would be the proper approach, given the data limitations on the U.S. sample.

3.2.6 PSE Used the Same Input Price Inflation Index Assumptions for the Entire Sample

We view using the same input price inflation indexes for the studied utility and the rest of the sample in a benchmarking study as the better approach. This issue is inconsequential, given that both PSE and PEG levelized input prices after the major inflation index differences between the PSE and PEG capital indexes occurred. Both PSE and PEG levelized the capital in 2012. We also note that PEG used a similar approach to PSE when it used U.S. inflation indexes in their recent Ontario Power Generation research.

3.2.7 Hydro One's OM&A Expenses Grow by the Proposed Revenue Escalation Formula (i.e., Inflation)

PSE escalated Hydro One's OM&A expenses in the forecasted period based on the proposed revenue escalation formula of I – X, where X = 0. PEG takes the exact same approach in their benchmarking research.

PEG believes that the company's I-X revenue escalation formula will not provide the company with enough revenue escalation; we deduce this because PEG also believes it is a "rosy scenario" for expenses to increase by only inflation during the Custom IR period. When requested to offer PEG's view, PEG refused to provide an opinion on the appropriate OM&A productivity factor in the revenue escalation formula.[41]

3.2.8 Four Other Items

PEG lists four other concerns it describes as "less important.".

1. PEG correctly states that PSE used Toronto values to levelize the Company's construction cost index. PEG used the same values in its research.[42] PSE used the headquarter city for

[40] See PEG's response to EB-2019-0082, L1, Tab 1, Schedule 7, part c.

[41] Exhibit L1, Tab 1, Schedule 8, part c.

[42] See PEG's response to EB-2019-0082, L1, Tab 1, Schedule 9, part a.

every utility in the sample, including Hydro One, when levelizing the capital asset price. This is the consistent approach and given that Hydro One serves many remote areas of Ontario, where capital prices could be higher than in Toronto, this is a good approximation of Hydro One's capital price levelization.[43]

2. PEG mentions that PSE applied the capital price levelization in the wrong year. Like the prior concern, PEG also levelized in the exact same year as PSE, which is 2012.[44] This is inconsequential to the result.

3. PEG discusses the 1.65 declining balance parameter used by PSE to formulate the transmission depreciation rate. This is, again, inconsequential to the benchmarking result. PEG used the same approach as PSE in their HOSSM research.

4. PEG states that PSE only used transmission plant in calculating the capital price and quantity trends, even though a material portion of assets are recorded as general plant. The approach that PSE undertook enables a consistent approach between Hydro One and the U.S. sample. In contrast, PEG's approach is not consistent and treats Hydro One differently than the rest of the sample, due to Hydro One's inability to break out transmission and general plant. We do not dwell on this inconsistency in our critiques of PEG in Section 2 because we believe this is an inconsequential inconsistency.

4 Reply to PEG's Plan Design Comments

In this section we provide a reply to some of PEG's plan design comments. We did not investigate the actual capital needs of the Company, and do not know if the proposed capital spending amounts are necessary. From a high-level perspective, what we do know of the capital spending plan is this: at the proposed capital spending levels, the company's total costs during the 2020-2022 period are 32% <u>below</u> the expected levels. This is PSE's result, and PEG's result would be close, if the two PEG flaws discussed in Section 2 are corrected.

This benchmarking result should not be ignored when contemplating whether the capital needs of the company are at the proper amounts. A finding of 32% below cost is a strong one and provides evidence that the company is producing cost savings relative to the industry, but also may need to increase spending for a time relative to the industry.

PEG recognized the need for capital spending in the electric transmission industry and how the industry has changed over time in work for the Edison Electric Institute (EEI), which is the US investor-owned electric utility industry's trade group. In a 2015 EEI paper that PEG authored (*Alternative Regulation for Emerging Utility Challenges),* on p. 47 PEG recognized the need for increased investments in the transmission industry to help tackle these emerging challenges in the utility industry. PEG wrote that investments in the power transmission industry are "urgently needed investments." However, PEG's suggestions on reducing the capital spending proposal of

[43] In PEG's response to EB-2019-0082, L1, Tab 1, Schedule 9, part b PEG suggests there is evidence the construction costs will be lower than Toronto. However, PEG cites indexes for a number of relatively large municipalities but ignore the fact that Hydro One serves many remote areas that likely increase construction costs.

[44] See PEG's response to EB-2019-0082, L1, Tab 1, Schedule 4, part e.

Hydro One and their suggestions on ways to markdown the utility's capital-related revenue appear to contradict the view that investments in the transmission sector are urgently needed.

On p. 43 of the PEG Report, PEG states that "it seems desirable to consider how to make Custom IR more mechanistic, incentivizing, and fair to customers while still ensuring that it is reasonably compensatory over time for efficient distributors." However, many of PEG's comments and suggestions seem to be contrary to that statement. PEG's suggestions include: adding a special stretch factor to the C factor calculation, materiality thresholds, raising of the X factor, underfunding in the last year of the plan term, and reducing the budget by a material amount. These proposed items would either reduce the mechanistic nature of the Custom IR plan, reduce incentives, or would not be reasonably compensatory for an efficient firm.

The introduction of markdowns in the form of a supplemental stretch factor and underfunding of the utility detracts from the ability of the company to retain reasonable compensation and set an incentive plan that is customized to the needs of the company. PEG is unaware of any U.S. multi-year rate plans that have approved a supplemental stretch factor on capital. In PEG's decades of work for utility clients PEG has never recommended such a stretch factor.[45] PSE is also unaware of any such plans other than the Hydro One Distribution decision which included a supplemental stretch factor of 0.15% on capital. In that case, both PSE and PEG found Hydro One Distribution's total costs were considerably higher than the benchmarks (PSE found Hydro One Distribution's score to be +22%). In this case, Hydro One is also proposing a "progressive productivity" component that is equivalent to a 0.14% stretch factor in 2021 and 0.33% in 2022. Further the implicit stretch factor, if the productivity factor is set at 0.0%, is found to be larger in the transmission industry by both consultants.

PEG's research indicates the industry's TFP over the 1996 to 2018 time period is declining by -0.44%. While we acknowledge that PEG has suggested the productivity factor be set at this TFP result, which means a negative productivity factor, that it is an unlikely outcome based on prior decisions. If the productivity factor is set at 0.0%, as it was in HOSSM, there is a stretch factor already embedded in that result. In fact, even based on PEG's longer time period, it is a stretch factor that already exceeds PEG's suggested supplemental stretch factor. Further, both PEG and PSE find that the recent years exhibit even more negative TFP trends. This makes the 0.0% productivity factor even more challenging for the company.

When formulating its suggestions on items like the S-factor, PEG did not recognize the Company's progressive productivity proposal within its application.[46] Based on company estimations, this proposal is equivalent to an additional 0.14% stretch factor in 2021, and a 0.33% stretch factor in 2022.[47] The 0.14% and 0.33% is equivalent to a stretch factor on the full revenue requirement and not just the capital portion. Now that PEG has corrected its S-factor to 0.31% after identifying errors in its calculation,[48] the company's progressive productivity proposal is nearly the same magnitude.

[45] See PEG's response to EB-2019-0082, L1, Tab 1, Schedule 2, part b and c.

[46] EB-2019-0082, L1, Tab 1, Schedule 13, part d.

[47] See Hydro One's response to JT 2.42.

[48] EB-2019-0082, L1, Tab 1, Schedule 13, part a.

In our view it would not be fair or compensatory to add a supplemental stretch factor on top of the large implicit stretch factor, the normal stretch factor, and the progressive productivity. This is especially true when the benchmarking results demonstrate the company's cost levels are considerably lower than expected.

PEG's construction of the supplemental stretch factor makes the productivity factor and stretch factor based on the total cost benchmarking results essentially irrelevant. If the productivity factor is increased by 0.1%, then the S-factor is lowered by that same 0.1%. Likewise, if a stretch factor of 0.15% is decided on, this would lower the S-factor by 0.15%. This neuters the incentive properties of total cost benchmarking and productivity analysis. PEG's suggestions set the markdown at the value of the capital depreciation amount multiplied by a markdown percentage. If the X-factor goes up, then S goes down by the same amount and vice versa. Further, this markdown is not based on capital needs or any evidence, it is a pre-set markdown regardless of needs and cost performance assessments. This pre-set markdown is not compensatory to an efficient firm. We do not see why a multi-year custom incentive regulation plan should be set equal to the ACM materiality threshold or how this justifies essentially eliminating key incentive components proposed by the Company such as the productivity factor and stretch factor.

We recommend that the Board not impose a supplemental stretch factor on capital in recognition of the following:

1. the benchmarking results provide strong evidence that Hydro One is efficient making it more difficult for the Company to achieve productivity savings relative to the industry;
2. the Company's plan already includes a progressive productivity component that essentially already acts as a supplemental stretch factor;
3. the already large implicit stretch factor of either 0.44% or 1.61% if the productivity factor is set at 0.0%; and
4. the presence of a S-factor based on the way PEG calculates it, negates the incentive properties of the productivity factor and stretch factor based on cost benchmarking results.

5 Concluding Remarks

PSE continues to recommend a productivity factor of 0.0% and a stretch factor of 0.0%, with no other supplemental stretch factors or systematic markdowns that are not connected to the capital needs of the Company. Both PEG and PSE find negative productivity in the transmission industry, and both firms find that a 0.0% productivity factor would already contain a substantial implicit stretch factor. Adding 2017 and 2018 to the sample provides further evidence of negative productivity trends, especially in the most recent years of the sample. With all of this, a 0.0% productivity factor is a difficult and challenging expectation for the company to meet and will likely exceed the productivity of the industry during the 2021 and 2022 years.

After updating the benchmarking dataset to 2018, PSE finds that Hydro One's total costs are 32.9% below benchmark expectations. This is extraordinary cost performance that should be recognized with a 0.0% stretch factor, especially considering Hydro One's proposed progressive productivity component. PEG has produced a model result that is unstable and inconsistent with its own research in the recent HOSSM case. It contains a clear bias against the recent and forecasted years for the entire sample, including Hydro One. When this bias is mitigated and PEG's modeling procedure corrected to what it used in HOSSM (or if PEG used the modern approach by using the OLS coefficients that do not require special coding, are transparent, cannot be improved upon, and do not require assumptions by the researcher), PEG's results would also indicate strong cost performance and a stretch factor of 0.0%.

Journal of Statistical Software

November 2017, Volume 82, Issue 3. *doi: 10.18637/jss.v082.i03*

Robust Standard Error Estimators for Panel Models: A Unifying Approach

Giovanni Millo
Group Insurance Research, Generali SpA
DEAMS, University of Trieste

Abstract

The different robust estimators for the standard errors of panel models used in applied econometric practice can all be written and computed as combinations of the same simple building blocks. A framework based on high-level wrapper functions for most common usage and basic computational elements to be combined at will, coupling user-friendliness with flexibility, is integrated in the **plm** package for panel data econometrics in R. Statistical motivation and computational approach are reviewed, and applied examples are provided.

Keywords: panel data, covariance matrix estimators, R.

1. Introduction

This paper is about computing estimators for the covariance matrix of parameters in a linear panel model, of the kind commonly used in applied practice to produce "robust" standard errors. Different estimators are usually preferred in one or the other branch of applied econometrics, from large microeconometric panels (Arellano 1987) to moderately-sized panel time series in macroeconomics (Driscoll and Kraay 1998) and large panels in finance (Petersen 2009; Cameron, Gelbach, and Miller 2011; Thompson 2011), up to pooled time series in political science (Beck and Katz 1995). Software implementations are in most cases to be found in one or the other commercial package, often as user-programmed additional routines; or sometimes GAUSS (Aptech Systems, Inc. 2006) or MATLAB (The MathWorks Inc. 2017) code is available. My work aims at bringing them all together under the common umbrella of the R environment (R Core Team 2017), once again the all-purpose statistical software.

From a software design viewpoint, I translate some results from the recent literature (Petersen 2009; Thompson 2011; Cameron *et al.* 2011) into a comprehensive computational framework, in turn integrated into the **plm** package for panel data econometrics (Croissant and Millo

2008). I describe a general expression for "clustering" estimators; then I review two-level clustering as a combination of simple clustering estimators and the extension to persistent effects by summation of lagged terms; lastly, I show how applying a weighting scheme to lagged covariance terms yields Driscoll and Kraay (1998)'s spatial correlation consistent (SCC) estimator (and, as a special case, that of Newey and West 1987).

From an application perspective, I extend the treatment of Petersen (2009) to double-clustering estimators plus time-persistent shocks as in Thompson (2011): a structure which, based on simulations in Petersen (2009), can be conjectured to successfully account for both individual effects and persistent idiosyncratic shocks. My approach also allows easy extension to a combination of effects which has not, to my knowledge, received attention in the literature yet: double-clustering as in Cameron *et al.* (2011) plus time-decaying correlation as in Driscoll and Kraay (1998). A practical example is given in Section 6.

One not-so-minor aim of this paper is to overcome sectoral barriers between different, if contiguous, disciplines: it is striking, for example, how few citations Driscoll and Kraay (1998) on the panel generalization of the Newey and West (1987) estimator gets in the finance literature, especially in those papers that advocate what is essentially an unweighted version of their SCC covariance. Also, Arellano (1987) and Froot (1989), in the different contexts of fixed effects panels with serial correlation and of industry-clustered financial data, independently developed what is computationally the same estimator (referred in the following as one-way clustering) first described by Liang and Zeger (1986). Cross-referencing between the different branches of statistical and econometric research is still uncommon in this subject, to the point that raising awareness might be useful.[1] From the point of view of political science, where panel – or time-series cross-section (TSCS) – data are an important methodological field, the functionality outlined here allows researchers to progress beyond the now-ubiquitous application of panel-corrected standard errors (PCSE, Beck and Katz 1995) to pooled specifications, along the lines of Wilson and Butler (2007): both comparing it with alternative strategies and possibly combining it with individual effects, in order to tackle the all-important, and often overlooked, issue of individual heterogeneity (Wilson and Butler 2007, Section 2.2).

In this sense, my work is meant to provide R users with a comprehensive set of modular tools: lower level components, conceptually corresponding to the statistical "objects" involved (see Zeileis 2006), and a higher-level set of "wrapper functions" corresponding to standard covariance estimators as they would be used in statistical packages: White heteroskedasticity-consistent, clustering, SCC and so on. Wrappers work by combining the same, few lower-level components in multiple ways in the spirit of the *Lego principle* of Hothorn, Hornik, Van De Wiel, and Zeileis (2006), with obvious benefits for both flexibility and code maintenance. This toolset should enable users to follow the work of Petersen (2009); Cameron *et al.* (2011); Thompson (2011) in detail, experimenting with settings and comparing estimates' magnitudes (see Petersen 2009) for specification and diagnostic purposes, at least as far as linear models in two panel dimensions are concerned.

Clustered standard errors for non-panel models are another field of application. For some time, there has been R code available for one- or two-way clustering in a linear model (see Arai 2009). This last has recently evolved into a package for multi-way clustering, **multiwayvcov**

[1]Also note that Fama and MacBeth (1973)'s covariance estimator popular in finance, actually first and foremost an estimator for the averages of the coefficients, is known in the econometrics literature as the Mean Group estimator of Pesaran and Smith (1995). See Ibragimov and Müller (2010) for a formal justification of the Fama-MacBeth method.

(Graham, Arai, and Hagströmer 2016); in turn, many of the features of the latter have been incorporated into the **sandwich** package by Berger, Graham, and Zeileis (2017). The **sandwich** package was the original, object-oriented implementation of sandwich estimators in R (Zeileis 2006) and provides the generic function `vcovHC`, panel methods for which are presented here. Nevertheless, up to two clustering dimensions all this functionality is effectively encompassed by that presented here, provided the data are treated like a *faux* panel specifying one or two indices. Moreover, integration within the **plm** package means that the estimators presented here can seamlessly interact with panel features like individual or time effects. By contrast, extending clustering to more than two dimensions in a panel context does not fit into the panel data infrastructure of package **plm** and is out of the scope of this paper.

When estimating regression models, R creates a model object which, besides estimation results, carries on a wealth of useful information, including the original data. Robust testing in R is done retrieving the necessary elements from the model object, using them to calculate a robust covariance matrix for coefficient estimates and then feeding the latter to the actual test function, which can be a t-test for significance, a Wald restriction test and so on. Therefore the approach to diagnostic testing is more flexible than with procedural languages, where diagnostics usually come with standard output. In our case, for example, one can obtain different estimates of the standard errors under various kinds of dependence without re-estimating the model, and present them compactly.

When appropriate, I will highlight some features of R that make it easy and effective to combine statistical objects; in particular, functions as arguments; modularity and components reusing; function application over arrays of arbitrary dimension; and in general object orientation, which ensures application of the right method with the appropriate defaults for the object at hand through standard syntax.

The paper is organized into three main bodies. The next two sections (Sections 2 and 3) review the statistical foundations of the methods and set the notation in terms of a few low-level components according to the Lego principle. Section 4 on the computational framework, arguably the heart of the paper, describes the statistical building blocks in terms of computational objects characterized by a few standard "switches", and their combinations in terms of user-friendly "wrapper" functions; then, in an object-oriented fashion, it discusses how and when it is (statistically) appropriate to apply the resulting user-level methods to '`plm`' objects estimated in different ways: by either (pooled) ordinary least squares (OLS), fixed effects (FE), random effects (RE), or by first-differencing methods (FD). The remainder of the paper (Sections 5 and 6) sets the new estimators in the context of the **plm** package and provides some examples of application.

The functionality described here is available in package **plm** since version 1.5-1 and the package is available from the Comprehensive R Archive Network (CRAN) at `https://CRAN.R-project.org/package=plm`.

2. Robust covariance estimators

In this section I briefly review the ideas behind robust covariance estimators of the *sandwich* type, in order to provide a basis for the subsequent treatment of their panel extension. The reader is referred to any econometrics textbook, e.g., Greene (2003) – on which this paragraph is based – for a formal treatment.

Consider a linear model $y = X\beta + \epsilon$ and the OLS estimator $\hat{\beta}_{OLS} = (X^\top X)^{-1}X^\top y$. If one is interested in making inference on β, then an estimate of $\mathsf{VAR}(\hat{\beta})$ is needed. If the error terms ϵ are independent and identically distributed, then the covariance matrix takes the familiar textbook form: $\mathsf{VAR}(\hat{\beta}) = \hat{\sigma}^2(X^\top X)^{-1}$, where $\hat{\sigma}^2$ is an estimate of the error variance. This case is synthetically dubbed *spherical errors*, and the relative formulation of $V(\hat{\beta}_{OLS})$ is often referred to, somewhat inappropriately, as "OLS covariance"[2].

Let us consider robust estimation in the context of the simple linear model outlined above. The problem at hand is to estimate the covariance matrix of the OLS estimator relaxing the assumptions of serial correlation and/or homoskedasticity without imposing any particular structure to the errors' variance or interdependence.

As the estimator of the OLS parameters' covariance matrix is

$$\hat{V} = \frac{1}{n}\left(\frac{X^\top X}{n}\right)^{-1}\left(\frac{1}{n}X^\top[\sigma^2\Omega]X\right)\left(\frac{X^\top X}{n}\right)^{-1}$$

in order to consistently estimate V it is not necessary to estimate all the $n(n+1)/2$ unknown elements in the Ω matrix but only the $K(K+1)/2$ ones in

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}\mathbf{x}_i\mathbf{x}_j^\top,$$

which may be called the *meat* of the sandwich, the two $\left(\frac{X^\top X}{n}\right)^{-1}$ being the *bread*. (From now on, we will concentrate exclusively on the meat, and we will dispose of the $1/n$ terms throughout.) All that is required are *pointwise consistent* estimates of the errors, which is satisfied by consistency of the estimator for β (see Greene 2003). In the heteroskedasticity case, correlation between different observations is ruled out and the *meat* reduces to

$$S_0 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\mathbf{x}_i\mathbf{x}_i^\top,$$

where the n unknown σ_i^2s can be substituted by e_i^2 (see White 1980). In the serial correlation case, the natural estimation counterpart would be

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}e_i e_j\mathbf{x}_i\mathbf{x}_j^\top,$$

but this structure proves too general to achieve convergence. Newey and West (1987) devise a (heteroskedasticity and) autocorrelation consistent estimator that works based on the assumption of correlation dying out as the distance between observations increases. The Newey-West HAC estimator for the *meat* takes that of White and adds a sum of covariances between the different residuals, smoothed out by a *kernel function* giving weights decreasing with distance:

[2]The reason is that OLS is "best linear unbiased" (BLUE) under sphericity; yet this is confusing because other covariance estimators can be more appropriate for $\hat{\beta}_{OLS}$ under different conditions. Notice that Thompson (2011) uses the same name referring to the case of heteroskedasticity but no dependence (here: *White*).

$$S_0 + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_l e_t e_{t-l} \left(\mathbf{x}_t \mathbf{x}_{t-l}^\top + \mathbf{x}_{t-l} \mathbf{x}_t^\top \right)$$

with w_l the weight from the kernel smoother, e.g., the Bartlett kernel function: $w_l = 1 - \frac{l}{L+1}$ (for a discussion of alternative kernels see Zeileis 2006). The lag l is usually truncated well below sample size: one popular rule of thumb is $L = n^{1/4}$ (see Greene 2003; Driscoll and Kraay 1998).

In the following I will consider the extensions of this framework for a panel data setting where, thanks to added dimensionality, various combinations of the two above structures will turn out to be able to accommodate very general types of dependence.

3. Clustering estimators in a panel setting

Let us now consider a *panel* (or *longitudinal*) setting where data are collected on two different dimensions: to fix ideas, let us think of n entities (individuals, firms, countries, …) which we here label *groups* and index by $i = 1, \ldots, n$ sampled at T points in *time*. The two dimensions are not fully symmetric as for the sake of practical relevance I have considered one dimension (time) having a natural ordering and the other having none. This setting is sufficiently general to describe the vast majority of applications; a symmetric extension would nevertheless be straightforward. Different choices of dimensions are possible and have been explored in the literature: e.g., Froot (1989), in the context of financial data, discusses sampling from "independent" industries in order to increase sample size, clustering within industry to account for dependence. Thus the two dimensions would be *industry* and a generic counter: clustering would be done according to industry, while meaningless across the "other" dimension. The model is then

$$y_{it} = X_{it} \beta + \epsilon_{it}.$$

For now I consider again the familiar OLS estimator $\hat{\beta}_{OLS}$, which in this setting is referred to as *pooled OLS* because it *pools* all observations together irrespective of their belonging to a given group (but see Section 4.4 for an extension to three other common panel estimators).

Clustering estimators work by extending the "sandwich" principle to panel data. Besides heteroskedasticity, the added dimensionality allows to obtain robustness against totally unrestricted timewise or cross-sectional correlation, provided this is along the "smaller" dimension. In the case of "large-N" (*wide*) panels, the big cross-sectional dimension allows robustness against serial correlation[3]; in "large-T" (*long*) panels, on the converse, robustness to cross-sectional correlation can be attained drawing on the large number of time periods observed. As a general rule, the estimator is asymptotic in the number of clusters: see Cameron *et al.* (2011, Section 2).

Imposing cross-sectional (serial) independence in fact restricts all covariances between observations belonging to different individuals (time periods) to zero, yielding an error covariance

[3]This is the case of the seminal contribution by Arellano (1987).

matrix with a block-diagonal structure: in the former case, $V(\epsilon) = I_n \otimes \Omega_i$, where

$$\Omega_i = \begin{bmatrix} \sigma^2_{i1} & \sigma_{i1,i2} & \dots & \dots & \sigma_{i1,iT} \\ \sigma_{i2,i1} & \sigma^2_{i2} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \sigma^2_{iT-1} & \sigma_{iT-1,iT} \\ \sigma_{iT,i1} & \dots & \dots & \sigma_{iT,iT-1} & \sigma^2_{iT} \end{bmatrix} \tag{1}$$

and the consistency relies on the cross-sectional dimension being "large enough" with respect to the number of free covariance parameters in the diagonal blocks. The other case is symmetric.

3.1. White-Arellano, or one-way clustering

White's heteroskedasticity-consistent covariance matrix[4] has been extended to clustered data by Liang and Zeger (1986) and to econometric panel data by Arellano (1987), who applied it in a fixed effects setting. Observations can be clustered by the cross-sectional (group) index which is arguably the most popular use of this estimator, and is appropriate in *large, short* panels because it is based on n-asymptotics; or by the time index, which is based on T-asymptotics and therefore appropriate for *long* panels. In the first case, the covariance estimator is robust against cross-sectional heteroskedasticity and also against serial correlation of arbitrary form. In the second case, symmetrically, against timewise heteroskedasticity and cross-sectional correlation. Arellano's original estimator, an instance of the first case, has the form:

$$V_{\text{White-Arellano}} = (X^\top X)^{-1} \sum_{i=1}^{n} X_i^\top u_i u_i^\top X_i (X^\top X)^{-1}. \tag{2}$$

It is of course still feasible to rule out serial correlation and compute an estimator that is robust to heteroskedasticity only, based on the following error structure:

$$\Omega_i = \begin{bmatrix} \sigma^2_{i1} & \dots & \dots & 0 \\ 0 & \sigma^2_{i2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & \sigma^2_{iT} \end{bmatrix} \tag{3}$$

in which case the original White estimator applies:

$$V_{\text{White}} = (X^\top X)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} u_{it}^2 \mathbf{x}_{it} \mathbf{x}_{it}^\top (X^\top X)^{-1}. \tag{4}$$

Some notation

Before discussing bidirectional extensions of this estimator, for the sake of clarity I will now define the "meat" of the two versions of the Arellano estimator, henceforth $V_{C.}$, with respect to

[4] See White (1980, 1984).

the clustering dimension: the original, group-clustered version, robust vs. heteroskedasticity and *serial* dependence, will be

$$V_{CX} = \sum_{i=1}^{n} X_i^\top u_i u_i^\top X_i, \tag{5}$$

while the time-clustered version, robust vs. heteroskedasticity and *cross-sectional* dependence, will be:

$$V_{CT} = \sum_{t=1}^{T} X_t^\top u_t u_t^\top X_t. \tag{6}$$

The standard White estimator on pooled data, which is symmetric w.r.t. clustering,

$$V_W = \sum_{t=1}^{T}\sum_{i=1}^{n} u_{it}^2 \mathbf{x}_{it}\mathbf{x}_{it}^\top \tag{7}$$

will be conveniently written as

$$V_W = \sum_{i=1}^{n} X_i^\top \mathrm{diag}(u_i^2) X_i = \sum_{t=1}^{T} X_t^\top \mathrm{diag}(u_t^2) X_t, \tag{8}$$

where $\mathrm{diag}(u^2)$ is a matrix with squares of elements of u on the diagonal and zeros elsewhere, so that all of these expressions share the common structure

$$V_{C\cdot} = \sum_{\cdot} X_\cdot^\top \mathsf{E}(u) X_\cdot \tag{9}$$

with $\mathsf{E}(u)$ an appropriate function of the residuals.

This symmetric representation will provide a good starting point for the extension to double clustering.

3.2. Double clustering

Some recent research in finance (Petersen 2009; Cameron *et al.* 2011; Thompson 2011) advocates double clustering, motivating it by the need to account for *persistent shocks* and at the same time for cross-sectional or spatial correlation.

This estimator combining both individual and time clustering relies on a combination of the asymptotics of each: the minimum number of clusters along the two dimensions must go to infinity: see, again, Cameron *et al.* (2011, Section 2). Apart from this, any dependence structure is allowed within each group *or* within each time period, while cross-serial correlations between observations belonging to different groups *and* time periods are ruled out.

The double-clustered estimator is easily calculated by summing up the group-clustering and the time-clustering ones, then subtracting the standard White estimator (referred to in Cameron *et al.* 2011 as *time-group-clustering*, in Thompson 2011 as *white0*) in order to avoid double-counting the error variances along the diagonal:

$$V_{CXT} = V_{CX} + V_{CT} - V_W. \tag{10}$$

In order to control for the effect of common shocks, Thompson (2011) proposes to add to the sum of covariances one more term, related to the covariances between observations from any group at different points in time. Given a maximum lag L, this will be the sum over $l = 1, \ldots, L$ of the following generic term:

$$V_{CT,l} = \sum_{t=1}^{T} X_t^\top u_t u_{t-l}^\top X_{t-l} \tag{11}$$

representing the covariance between pairs of observations from any group distanced l periods in time, summed with its transpose. As the correlation between observations belonging to the *same* group at different points in time has already been captured by the group-clustering term, to avoid double counting one must subtract the within-groups part:

$$V_{W,l} = \sum_{t=1}^{T} \sum_{i=1}^{n} [x_{it} u_{it} u_{i,t-l}^\top x_{i,t-l}^\top] \tag{12}$$

again summed with its transpose, for each l. The resulting estimator

$$V_{CXT,L} = V_{CX} + V_{CT} - V_W + \sum_{l=1}^{L} [V_{CT,l} + V_{CT,l}^\top] - \sum_{l=1}^{L} [V_{W,l} + V_{W,l}^\top] \tag{13}$$

is robust to cross-sectional and timewise correlation inside, respectively, time periods and groups *and* to the cross-serial correlation between observations belonging to different groups, up to the Lth lag. It also resembles another well-known estimator from the econometric literature: the Newey-West nonparametric estimator, the difference being that instead of adding a (possibly truncated) sum of unweighted lag terms, the latter downweighs the correlation between "distant" terms through a kernel smoother function. Kernel-smoothed estimators will be the subject of the next section.

3.3. Kernel-based smoothing

As cited above, in a time series context Newey and West (1987) proposed an estimator that is robust to serial correlation as well as to heteroskedasticity. This estimator, based on the hypothesis of the serial correlation dying out "quickly enough", takes into account the covariance between units by: weighting it through a kernel smoother function giving less weight as they get more distant; and adding it to the standard White estimator.

Driscoll and Kraay's "SCC"

Driscoll and Kraay (1998) adapted the Newey-West estimator to a panel time series context, where not only serial correlation between residuals from the same individual in different time periods is taken into account, but also cross-serial correlation between different individuals in different times and, within the same period, cross-sectional correlation (see also Arellano 2003, p. 19).

The Driscoll and Kraay estimator, labeled SCC (as in "spatial correlation consistent"), is defined as the time-clustering version of Arellano plus a sum of lagged covariance terms, weighted by a distance-decreasing kernel function w_l:

$$\begin{aligned} V_{SCC,L} &= V_{CT} + \textstyle\sum_{l=1}^{L} w_l [\sum_{t=1}^{T} X_t^\top u_t u_{t-l}^\top X_{t-l} + \sum_{t=1}^{T} [X_t^\top u_t u_{t-l}^\top X_{t-l}]^\top] \\ &= V_{CT} + \textstyle\sum_{l=1}^{L} w_l [V_{CT,l} + V_{CT,l}^\top]. \end{aligned} \tag{14}$$

The "scc" covariance estimator requires the data to be a mixing sequence, i.e., roughly speaking, to have serial and cross-serial dependence dying out quickly enough with the T dimension, which is therefore supposed to be fairly large: Driscoll and Kraay (1998), based on Monte Carlo simulation, put the practical minimum at $T > 20 - 25$; the n dimension is irrelevant in this respect and is allowed to grow at any rate relative to T.

Panel Newey-West

By restricting the cross-sectional and cross-serial correlation to zero one gets a "pure" panel version of the original Newey-West estimator, as discussed, e.g., in Petersen (2009):

$$\begin{aligned} V_{NW,L} &= V_W + \textstyle\sum_{l=1}^{L} w_l [\sum_{t=1}^{T} \sum_{i=1}^{n} [\mathbf{x}_{it} u_{it} u_{i,t-l}^{\top} \mathbf{x}_{i,t-l}^{\top}] + \sum_{t=1}^{T} [\sum_{i=1}^{n} [\mathbf{x}_{it} u_{it} u_{i,t-l}^{\top} \mathbf{x}_{i,t-l}^{\top}]^{\top}] \\ &= V_W + \textstyle\sum_{l=1}^{L} w_l [V_{W,l} + V_{W,l}^{\top}]. \end{aligned} \tag{15}$$

As is apparent from Equation 14, if the maximum lag order is set to 0 (no serial or cross-serial dependence is allowed) the SCC estimator reverts to the cross-section version (time-clustering) of the Arellano estimator V_{CT}. On the other hand, if the cross-serial terms are all unweighted (i.e., if $w_l = 1\, \forall l$) then $V_{SCC,L|w=1} = V_{CT,L}$.

3.4. Unconditional estimators

Unconditional covariance estimators are based on the assumption of no error correlation in time (cross-section) and of an unrestricted but invariant correlation structure inside every cross-section (time period). They are popular in contexts characterized by relatively small samples, with prevalence of the time dimension.

Beck and Katz PCSE

Beck and Katz (1995), in the context of political science models with moderate time and cross-sectional dimensions, introduced the so-called panel corrected standard errors (PCSE), an estimator with good small-sample properties which, in the original time-clustering setting, is robust against cross-sectional heteroskedasticity and correlation.

In this framework and with reference to Equation 9, the "pcse" covariance is defined in terms of the $E_i = E\ \forall i$ function of the residuals as:

$$E = \frac{\sum_n \hat{e}_n \hat{e}_n^{\top}}{N}.$$

A sufficient, although not necessary condition for consistency of the "pcse" estimator (Beck and Katz 1996) is that the covariance matrix of the errors in every group be the same: $\Omega = \Sigma \otimes I_T$, with

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \dots & \sigma_{1,T} \\ \sigma_{2,1} & \sigma_2^2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \sigma_{T-1}^2 & \sigma_{T-1,T} \\ \sigma_{T,1} & \dots & \dots & \sigma_{T,T-1} & \sigma_T^2. \end{bmatrix} \tag{16}$$

A possible further restriction is to assume correlation away imposing that Σ be diagonal, thus restricting the estimator to unconditional intragroup heteroskedasticity.

4. Computational framework

Generalizing the computational problem at hand and dividing it into modules is necessary for writing software that be both full-featured and easy to read and to maintain. In this section I show a generic formulation capable of generating all the estimators considered up to now; in the following I will consider a small-sample correction module. These building blocks will allow to construct a very general covariance estimating function whose usage in various testing environments will then be discussed in the light of object-oriented econometric computing.

Two defining features of R as a programming language are object-orientation and functional nature. In this sense, according to the object-oriented nature of R, in the next paragraph I will formulate a general computing module, the *software counterpart* of the *statistical objects* V_W, V_{CX}, V_{CT}, $V_{W,l}$, $V_{CX,l}$ and $V_{CT,l}$ which are in turn the building blocks for any of the estimators considered here. In turn, according to the functional nature of R, the computing module will be formulated as a function of: a (panel-indexed) vector of errors; an integer lag order; and lastly of a function to be applied to the error vector, taking the lag order as an argument. The ability of R to treat functions as a data type will make the translation of this formalization into software seamless.

4.1. A general, computing-oriented formulation

The most general formulation of the comprehensive estimator can be written as a kernel-weighted version of Formula 3 in Thompson (2011):

$$V_{CXT,L|w} = V_{CT} + \sum_{l=1}^{L} w_l[V_{CT,l} + V_{CT,l}^\top] + V_{CX} - V_W - \sum_{l=1}^{L} w_l[V_{W,l} + V_{W,l}^\top]. \tag{17}$$

In turn, all building blocks for Equation 17 can be generated by combining a clustering dimension (n or t), a lag order l and a function of the errors f. Starting from the general formulation:

$$V_g(t, l, f) = \sum_{t=1}^{T} X_t^\top f(u_t, u_{t-l}) X_{t-l} \tag{18}$$

inserting the outer product function and setting the lag to zero (so that $f(u_t, u_{t-l}) = u_t u_t^\top$) we get the time- (group-)clustering estimator

$$V_{CT} = V_g(t, 0, f = u_t u_t^\top) \tag{19}$$

and for the "White" terms on the diagonal, with the dot denoting indifferently n or t as clustering dimension,

$$V_W = V_g(\cdot, 0, f = \mathrm{diag}(u_\cdot^2)), \tag{20}$$

while for the cross-serial terms

$$V_{CT,l} = V_g(t, l, f = u_t u_{t-l}^\top) \tag{21}$$

Label	Notation
White heteroskedastic	V_W
Group clustering	V_{CX}
Time clustering	V_{CT}
Double clustering	$V_{CXT} = V_{CX} + V_{CT} - V_W$
Time clustering + shocks	$V_{CT,L} = V_{CT} + \sum_{l=1}^{L}[V_{CT,l} + V_{CT,l}^{\top}]$
Panel Newey-West	$V_{NW,L} = V_W + \sum_{l=1}^{L} w_l[V_{W,l} + V_{W,l}^{\top}]$
Driscoll and Kraay's SCC	$V_{SCC,L} = V_{CT} + \sum_{l=1}^{L} w_l[V_{CT,l} + V_{CT,l}^{\top}]$
Double-clustering + shocks	$V_{CXT,L} = V_{CT} + \sum_{l=1}^{L}[V_{CT,l} + V_{CT,l}^{\top}] + V_{CX}$ $-V_W - \sum_{l=1}^{L}[V_{W,l} + V_{W,l}^{\top}]$ $= V_{CT,L} + V_{CX} - V_{NW,L\|w=1}$

Table 1: Covariance structures as combinations of the basic building blocks.

and for the purely autoregressive ones:

$$V_{W,l} = V_g(t, l, f = \mathrm{diag}(u_t \times u_{t-l})) \tag{22}$$

(where $\times$ is the element-by-element product) so that by a (possibly weighted) combination of the above we can get all relevant estimators: see Table 1.

As observed, the SCC estimator differs from the (one-way) time-shocks-robust version of the double-clustering à la CGM only by the distance-decaying weighting of the covariances between different periods, so that $V_{CT,L} = V_{SCC,L|w=1}$.

Obviously, as there is no natural univariate ordering between individuals, a full generalization encompassing both the double-clustering and a two-way SCC estimator does not seem sensible. For the same reason, while the software components allow fully symmetric operation, i.e., it would be practically feasible to compute a group-clustering version of $V_{SCC,L}$ or $V_{CX,L}$, this is devoid of sense from a statistical viewpoint because the notion of a linear, unidimensional spatial lag is generally meaningless[5].

4.2. Unconditional estimation in the general framework

Unconditional estimators can also be computed from the general formulation by precalculating the unconditional error covariance $E = \frac{\sum_{t=1}^{T} u_t u_t^{\top}}{T}$ and substituting it inside the generic Equation 17 as a constant matrix:

$$V_{UT} = V_g(t, 0, f = E). \tag{23}$$

One noteworthy feature of R is the ability to treat functions as first-class objects (R Core Team 2017), which means they are just another, although very special, data type and can be fed to another function as argument. So a function might indifferently take as argument a function or a precalculated matrix, which is the case here[6].

[5] Spatial lags, where applicable, are defined in a completely different way based on two-dimensional proximity matrices (see Anselin 1988). One very special case of linear spatial arrangement allowing for a simple definition of lag is Chudik, Pesaran, and Tosetti (2011)'s circular world, where each observation has one neighbor to the left and one to the right. Yet in that setting dependence would have to consider both directions, while serial dependence only originates from the past.

[6] Another example of use of this powerful feature for passing a covariance matrix or the function calculating it is in Zeileis (2006).

4.3. Unbalanced panels

Unbalancedness is one of the major computational nuisances in panel data econometrics. In the case at hand, the problem is to compute the generic formula in Equation 17 taking heed that unbalanced samples will have incomplete groups (time periods) for some t (i). As the ultimate goal of estimation of the *meat* is to average the $k \times k$ matrix products $X_t^\top f(u_t, u_{t-l})X$ over time periods (or, symmetrically, groups), missing data will give rise to empty positions in X_t and, correspondingly, in $f(u_t, u_{t-l})$.

Fortunately, R has two particularly useful features for treating data in a "generic" way, as independent as possible from dimensions: `list` objects and the `apply` family of functions.

In general R makes it relatively easy to deal with unbalanced data through the use of structures like `list`s, very flexible containers which can hold e.g., matrices of different dimensions and on which operators (and, more generally, *functions* of any kind) can be *applied*. The `apply` family has members for working member-by-member on `lists` (`lapply`), subgroup by subgroup on *ragged arrays* (`tapply`) where (possibly heterogeneous) subgroups are defined by a grouping variable, or dimension by dimension on arrays, which is the original use of `apply`. One notable advantage of this operator is that it can work on arbitrary subsets of the array's dimensions, provided the function to be applied is compatible.[7]

If a function is *applied* that allows discarding of `NA` values, one can easily get consistent averaging over multidimensional arrays: in this case, an average of t $k \times k$ bidimensional matrices of uneven dimensions.

An example will clarify things. Let us take an array of three 3×3 matrices with a missing value, and average over the third dimension. By default, missing values propagate throughout operations:

```
R> a <- array(1, dim = c(3, 3, 3))
R> a[1, 1, 1] <- NA
R> apply(a, 1:2, mean)

     [,1] [,2] [,3]
[1,]   NA    1    1
[2,]    1    1    1
[3,]    1    1    1
```

but the default behavior can be overridden forcing discarding of `NA`s:

```
R> apply(a, 1:2, mean, na.rm = TRUE)

     [,1] [,2] [,3]
[1,]    1    1    1
[2,]    1    1    1
[3,]    1    1    1
```

[7]Notable examples of the usefulness of `lists` and *ragged arrays* for unbalanced panel data econometrics are, respectively, one-by-one inversion of lists of submatrices in general GLS calculations and time- (group-) demeaning of data based on grouping indices; both in package **plm**.

In the latter case, the resulting 3×3 matrix will contain all averages computed on the correct number of items (i.e., for the $[1, 1]$ position, $(1 + 1)/2$).

Analogously, in our case it will be convenient to make use of standard tridimensional arrays making a $k \times k \times t$ matrix – basically a "pile" of $X_t^\top f(u_t, u_{t-l})X$ terms – and then *applying* the `mean` function over the third dimension, obtaining an appropriate calculation of Equation 17 as a result. In fact, for every value of t every product involving a missing element will produce an `NA` in the relative $k \times k$ matrix; but then averaging over the T dimension will discard `NA`s and apply the correct denominator.[8]

The same goes for the estimation of the unconditional covariance in Beck and Katz type estimators. This feature, which has been unavailable for unbalanced panels for a while and then has been twice mentioned in the literature as a potentially complicated computational problem (Franzese 1996; Bailey and Katz 2011) is solved nicely and without effort in R by applying (sic!) the above method, which acts as advocated by Franzese (1996), averaging elements in the unconditional covariance matrix on the correct number of observations[9].

4.4. Application to FE, RE and FD models: The demeaning framework

From a software viewpoint, the methods provided here can be transparently applied either to pooled OLS or to any other model represented by a '`plm`' object. As usual, what is computationally feasible is not necessarily sound from a statistical viewpoint.

The application of the above estimators to pooled data is always warranted, subject to the relevant assumptions mentioned before. In some, but not all cases, these can also be applied to random or fixed effects panel models, or models estimated on first-differenced data. The general idea is to use both the covariates and residuals from the transformed (partially or totally demeaned, first differenced) model used in estimation.

In all of these cases the estimator is computed as OLS on transformed data, e.g., in the fixed effects case $\hat{\beta}_{FE} = (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top \tilde{y}$ with $\tilde{y}_{it} = y_{it} - y_{i.}$ and $\tilde{x}_{jit} = x_{jit} - x_{ji.}$ for each $\mathbf{x}_j$ in X; while in the random effects case this time-demeaning is partial and $\tilde{y}_{it} = y_{it} - \theta y_{i.}$ with $0 < \theta < 1$. Sandwich estimators can then be computed by applying the usual formula to the transformed data and residuals $\tilde{u} = \tilde{y} - \tilde{X}\hat{\beta}$: see Arellano (1987) and Wooldridge (2002, Section 10.59) for the fixed effects case, Wooldridge (2002, Chapter 10) in general.

In the following I discuss when it is appropriate to apply clustering estimators to the residuals of demeaned or first-differenced models.

Fixed effects

The fixed effects estimator requires particular caution. In fact, under the hypothesis of spherical errors in the original model, the time-demeaning of data induces a serial correlation $\mathsf{COR}(u_{it}, u_{i,t-1}) = -1/(T-1)$ in the demeaned residuals (see Wooldridge 2002, p. 275).

The White-Arellano estimator has originally been devised for this case. By way of symmetry, it can be used for time-clustered data with time fixed effects. The combination of group-

[8] Of course the most delicate programming issue becomes correct handling of the positions of incomplete u_t subvectors and X_t submatrices in the relevant tth "layer" of the tridimensional array.

[9] This estimation method, based on all available covariances between two given observations, corresponds to the `pairwise` option in the **pcse** function and package (Bailey and Katz 2011); it must be noted, though, that the default option there (`casewise`) is to use a balanced subset of the data.

clustering with time fixed effects and the reverse seems inappropriate because of the serial (cross-sectional) correlation induced by the time- (cross-sectional-) demeaning.

By analogy, the Newey-West type estimators can be safely applied to models with individual fixed effects (for an application, see Golden and Picci 2008), while the time and two-way cases require caution.

Random effects

In the random effects case, as Wooldridge (2002) notes, the quasi-time demeaning procedure removes the random effects reducing the model on transformed data to a pooled regression, thus preserving the properties of the White-type estimators.

By extension of this line of reasoning, all above estimators seem to be applicable to the demeaned data of a random effects model, provided *the transformed errors* meet the relevant assumptions.

First-differences

First-differencing, like fixed effects estimation, removes time-invariant effects. Roughly speaking, the choice between the two rests on the properties of the error term: if it is assumed to be well-behaved in the original data, then FE is the most efficient estimator and is to be preferred; if on the contrary the original errors are believed to behave as a random walk, then first-differencing the data will yield stationary and uncorrelated errors, and is therefore advisable (see Wooldridge 2002, p. 281). Given this, FD estimation is nothing else than OLS on differenced data, and the usual clustering formula applies (see Wooldridge 2002, p. 282). As in the RE case, the statistical properties of the various covariance estimators will depend on whether *the transformed errors* meet the relevant assumptions.

From the viewpoint of software implementation, the application to fixed or random effects and to first-difference models is greatly simplified by the availability in **plm** of a comprehensive data transformation infrastructure, allowing to easily extract the original data from the model object and apply the relevant transformation (see Croissant and Millo 2008).

4.5. Small-sample corrections

Two kinds of small-sample corrections are implemented: corrections for a small number of observations, derived from the work of MacKinnon and White (1985) and summarized in Zeileis (2006), and corrections for a small number of clusters, described in Cameron *et al.* (2011, p. 8).

All work by multiplying each residual by the square root of the appropriate correction factor $\sqrt{c}$, so that the various squares and cross-products of residuals are correctly multiplied by c while the correction can work at vector level, separating the small-sample-correction module from the other logical steps of computation. The cluster-level correction in turn works at single-clustering level, according to the relevant numerosity parameters, as suggested in Cameron *et al.* (2011): therefore small-sample cluster-level corrections for different clustering dimensions are seamlessly combined.

In all these cases $c > 0$, and $c \to 1$ as the total number of observations or, in the latter case, the number of clusters diverge.

5. R implementation

In this section I will first put the covariance estimators in the context of the **plm** package for panel data econometrics and provide a minimal background on robust restriction testing through interoperability between testing functions and covariance estimators. Then I will describe how the new approach detailed in this paper has been implemented, substituting the existing procedures in the simpler cases while extending functionality to the more complex ones. Lastly I will provide some applied examples to illustrate usage.

5.1. plm and generic sandwich estimators

Robust covariance estimators à la White or à la Newey and West for different kinds of regression models are available in package **sandwich** (Zeileis and Lumley 2017) under form of appropriate methods for the generic functions `vcovHC` and `vcovHAC` (Zeileis 2004, 2006). These are designed for data sampled along one dimension, therefore they cannot generally be used for panel data; yet they provide a uniform and flexible software approach which has become standard in the R environment. The procedures described in this paper have therefore been designed to be syntactically compliant with them.

plm (Croissant and Millo 2008) is an R package for panel data econometrics in which an S3 method for '`plm`' objects for the generic function `vcovHC` has long been available, allowing to apply sandwich estimators to panel models in a way that is natural for users of the **sandwich** package. In fact, despite the different structure "under the hood", the user will, e.g., specify a robust covariance for the diagnostics table of a panel model in the same way she would for a linear or a generalized linear model, the object-orientation features of R taking care that the right statistical procedure be applied to the model object at hand. What will change, though, are the defaults: the `vcovHC` method for '`lm`' objects defaults to the original White estimator, while the `vcovHC` method for '`plm`' objects to clustering by groups, both the most obvious choices for the object at hand.

As an example, Munnell (1990) specifies a Cobb-Douglas production function that relates the gross social product (gsp) of a given US state to the input of public capital (pcap), private capital (pc), labor (emp) and state unemployment rate (unemp) added to capture business cycle effects. Considering this model, whose dataset is a built-in example in **plm**,

```
R> library("plm")
R> data("Produc", package = "plm")
R> fm <- log(gsp) ~ log(pcap) + log(pc) + log(emp) + unemp
```

and the function `coeftest` from package **lmtest** (Zeileis and Hothorn 2002) which produces a compact coefficients table allowing for a flexible choice of the covariance matrix, I calculate the "robust" diagnostic table for two statistically equivalent models: OLS by `lm`

```
R> lmmod <- lm(fm, Produc)
R> library("lmtest")
R> library("sandwich")
R> coeftest(lmmod, vcov = vcovHC)

t test of coefficients:
```

```
              Estimate Std. Error t value  Pr(>|t|)    
(Intercept)  1.6433023  0.0716070 22.9489 < 2.2e-16 ***
log(pcap)    0.1550070  0.0186973  8.2903 4.668e-16 ***
log(pc)      0.3091902  0.0126283 24.4839 < 2.2e-16 ***
log(emp)     0.5939349  0.0197887 30.0139 < 2.2e-16 ***
unemp       -0.0067330  0.0013501 -4.9872 7.497e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

and pooled OLS by `plm`

```
R> plmmod <- plm(fm, Produc, model = "pooling")
R> coeftest(plmmod, vcov = vcovHC)

t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)    
(Intercept)  1.6433023  0.2441821  6.7298 3.211e-11 ***
log(pcap)    0.1550070  0.0601195  2.5783   0.01010 *  
log(pc)      0.3091902  0.0462297  6.6881 4.209e-11 ***
log(emp)     0.5939349  0.0686061  8.6572 < 2.2e-16 ***
unemp       -0.0067330  0.0030904 -2.1787   0.02964 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As can be seen, the estimated SEs will turn out different as the types of the model objects to be tested are different, unless one overrides the defaults: here specifying the `method` as `"white1"` and the small sample correction as `"HC3"` will replicate the `lm` results:

```
R> coeftest(plmmod,
+    vcov = function(x) vcovHC(x, method = "white1", type = "HC3"))

t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)    
(Intercept)  1.6433023  0.0716070 22.9489 < 2.2e-16 ***
log(pcap)    0.1550070  0.0186973  8.2903 4.668e-16 ***
log(pc)      0.3091902  0.0126283 24.4839 < 2.2e-16 ***
log(emp)     0.5939349  0.0197887 30.0139 < 2.2e-16 ***
unemp       -0.0067330  0.0013501 -4.9872 7.497e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As observed, these features have long been present in **plm**, but limited to one-way clustering (see Croissant and Millo 2008, Section 6.7); one-way SCC and the unconditional Beck and Katz (BK) estimator have also been added at a later stage, each one with its own infrastructure. With the exception of BK, this functionality has now been replaced by a combination of a general parameter covariance estimator as in Equation 17 and specific wrappers, replicating its different particularizations for the most common forms of usage.

5.2. The new modular framework

In this section I show how to use the basic "building block": the general estimator in Equation 17. This is unlikely to be much used in practice but it is left available at user level both for educational use and to possibly allow combinations not implemented in the higher-level wrappers. Then I show what is probably going to be the preferred option for practicing econometricians, that is the higher-level wrappers combining different particularizations of the general estimator to obtain one- or two-way clustering or kernel-weighted estimators à la White, Arellano, CGM, NW or DK. Lastly I show how to easily define custom combinations of the above to estimate more complicated covariance structures.

The general parameter covariance estimator has been implemented in R in a function `vcovG` which is the software counterpart to Equation 18 and can be used for calculating $V_{W,l}$, $V_{CT,l}$ or $V_{CX,l}$. This is visible at the user level and can be used as such, leaving the default lag at 0, to calculate V_W, V_{CT} or V_{CX}. According to the formalization in Equation 18, besides a '`plm`' object and a small-sample correction, it takes as arguments a clustering dimension (`cluster`), a function of the errors corresponding to $\mathsf{E}(u)$ in Equation 9 (`inner`) and a lag order. The `inner` argument accepts either one of two strings `"cluster"` or `"white"`, specifying respectively $\mathsf{E}(u) = uu^\top$ and $\mathsf{E}(u) = \mathrm{diag}(u^\top u)$, or a user-supplied function.

Next, I calculate the Arellano estimator V_{CX} by specifying `"group"` as the clustering dimension, `"cluster"` as the inner function and 0 as the lag order:

```
R> vcovG(plmmod, cluster = "group", inner = "cluster", l = 0)

              (Intercept)     log(pcap)       log(pc)      log(emp)
(Intercept)  0.0596248904 -9.637916e-03 -0.0068911857  0.0148866870
log(pcap)   -0.0096379163  3.614354e-03 -0.0002956929 -0.0031157168
log(pc)     -0.0068911857 -2.956929e-04  0.0021371841 -0.0017597732
log(emp)     0.0148866870 -3.115717e-03 -0.0017597732  0.0047067982
unemp        0.0003700792 -8.058266e-05 -0.0000586966  0.0001366349
                    unemp
(Intercept)  3.700792e-04
log(pcap)   -8.058266e-05
log(pc)     -5.869660e-05
log(emp)     1.366349e-04
unemp        9.550671e-06
attr(,"cluster")
[1] "group"
```

For the convenience of the user, a wrapper function `vcovHC` is provided which reproduces the syntax and results of the stand-alone version already available in **plm**, thus ensuring both retrocompatibility with **plm** and naming consistency with the **sandwich** package. Thus, the following statement reproduces the same output as above (suppressed) in a more intuitive way:

```
R> vcovHC(plmmod)
```

Higher-level functions are needed, and provided, in order to produce the double-clustering and kernel-smoothing estimators by (possibly weighted) sums of the former terms. The general tool in this respect, in turn based on `vcovG`, is `vcovSCC`, which computes weighted sums of $V_{\cdot,l}$ according to a weighting function which is by default the Bartlett kernel. Again, this function is available at user level and the default values will yield the Driscoll and Kraay estimator, $V_{SCC,L}$:

```
R> vcovSCC(plmmod)

              (Intercept)     log(pcap)        log(pc)      log(emp)
(Intercept)  0.0226046609 -5.514511e-03 -6.334497e-04  5.759358e-03
log(pcap)   -0.0055145106  1.367029e-03  1.319429e-04 -1.402905e-03
log(pc)     -0.0006334497  1.319429e-04  5.843328e-05 -1.862888e-04
log(emp)     0.0057593584 -1.402905e-03 -1.862888e-04  1.497875e-03
unemp       -0.0003377024  8.428261e-05  3.257782e-06 -8.034358e-05
                    unemp
(Intercept) -3.377024e-04
log(pcap)    8.428261e-05
log(pc)      3.257782e-06
log(emp)    -8.034358e-05
unemp        6.445790e-06
attr(,"cluster")
[1] "time"
```

No weighting (equivalent to passing the constant 1 as the weighting function: `wj = 1`) will produce the building blocks for double-clustering, according to Equation 13, so that V_{CXT} could be easily obtained defining it at user level as:[10]

```
R> myvcovDC <- function(x, ...) {
+    Vcx <- vcovHC(x, cluster = "group", method = "arellano", ...)
+    Vct <- vcovHC(x, cluster = "time", method = "arellano", ...)
+    Vw <- vcovHC(x, method = "white1", ...)
+    return(Vcx + Vct - Vw)
+  }
R> myvcovDC(plmmod)

              (Intercept)     log(pcap)        log(pc)      log(emp)
(Intercept)  0.0635274416 -1.087953e-02 -0.0067108330  0.0159466020
log(pcap)   -0.0108795286  3.809110e-03 -0.0002102193 -0.0033786244
log(pc)     -0.0067108330 -2.102193e-04  0.0020211433 -0.0017355810
log(emp)     0.0159466020 -3.378624e-03 -0.0017355810  0.0049283961
```

[10]Notice the use of the prefix “my” to indicate that this function has been defined by the user in this session, as opposed to built-in functions. This is done only for the sake of clarity, as R leaves complete naming freedom to the user; yet adhering to naming conventions of some sort is advisable in order to avoid inadvertently replacing built-in functions.

```
unemp        0.0002236813 -4.386756e-05 -0.0000544364  0.0000986291
                    unemp
(Intercept)  2.236813e-04
log(pcap)   -4.386756e-05
log(pc)     -5.443640e-05
log(emp)     9.862910e-05
unemp        1.108906e-05
attr(,"cluster")
[1] "group"
```

Again, convenience wrappers are provided to make usage more intuitive: `vcovNW` computes the panel Newey-West estimator $V_{NW,L}$ (output omitted); `vcovDC` the double-clustering one V_{CXT}, which is constructed not unlike `myvcovDC` from the example above, and gives exactly the same output (suppressed):

```
R> vcovDC(plmmod)
```

More complicated structures allowing for two-way clustering and error persistence in the sense of Thompson (2011) are easily obtained by combination, the same way as illustrated above, following the lines of Section 4.1. Below the case of double-clustering plus four periods of persistent (unweighted) shocks à la Thompson (2011) (notice that the weighting function `wj` has been defined as the constant 1 but must still be a function of two arguments):

```
R> myvcovDCS <- function(x, maxlag = NULL, ...) {
+    w1 <- function(j, maxlag) 1
+    VsccL.1 <- vcovSCC(x, maxlag = maxlag, wj = w1, ...)
+    Vcx <- vcovHC(x, cluster = "group", method = "arellano", ...)
+    VnwL.1 <- vcovSCC(x, maxlag = maxlag, inner = "white", wj = w1, ...)
+    return(VsccL.1 + Vcx - VnwL.1)
+  }
R> myvcovDCS(plmmod, maxlag = 4)
```

```
              (Intercept)     log(pcap)       log(pc)      log(emp)
(Intercept)  0.0766973526 -0.0160969792 -4.713237e-03  0.0191602519
log(pcap)   -0.0160969792  0.0043713347  2.332514e-04 -0.0042963693
log(pc)     -0.0047132370  0.0002332514  1.066283e-03 -0.0012435555
log(emp)     0.0191602519 -0.0042963693 -1.243556e-03  0.0052481667
unemp       -0.0006069241  0.0001587212 -9.439635e-06 -0.0001351121
                    unemp
(Intercept) -6.069241e-04
log(pcap)    1.587212e-04
log(pc)     -9.439635e-06
log(emp)    -1.351121e-04
unemp        1.403075e-05
attr(,"cluster")
[1] "time"
```

6. Applied examples

In the following applied examples, I will present the complete array of standard error estimates for each estimator in Table 1. A complete array of methods is presented for the sake of illustration; nevertheless one must keep in mind that the sample size and the number of clusters in either cross-section or time might prove inadequate for some estimators, as reported in the reference papers (see in particular Thompson 2011). The examples below must therefore be seen as examples of computational feasibility, not of statistical soundness of each method.

In fact, even limiting to those methods that are not at odds with the given sample size, the strategy of computing standard errors in all potentially sensible ways and taking the most conservative ones does indeed reduce type I error but at the same time decreases the power of the significance test.[11]

Another purpose of this section is to illustrate some ways to efficiently perform such multiple comparisons through some features of R. Looping on a vector of functions is a useful consequence of R treating functions as a data type. For the sake of clarity, let us predefine some functions for calculating the different covariance estimators in Section 4.1 according to the names reported there and with the appropriate parameters (leaving the maximum lag calculation at its default value of $L = T^{\frac{1}{4}}$):

```
R> Vw <- function(x) vcovHC(x, method = "white1")
R> Vcx <- function(x) vcovHC(x, cluster = "group", method = "arellano")
R> Vct <- function(x) vcovHC(x, cluster = "time", method = "arellano")
R> Vcxt <- function(x) Vcx(x) + Vct(x) - Vw(x)
R> Vct.L <- function(x) vcovSCC(x, wj = function(j, maxlag) 1)
R> Vnw.L <- function(x) vcovNW(x)
R> Vscc.L <- function(x) vcovSCC(x)
R> Vcxt.L <- function(x)
+    Vct.L(x) + Vcx(x) - vcovNW(x, wj = function(j, maxlag) 1)
```

then build up a vector of functions on which to loop:

```
R> vcovs <- c(vcov, Vw, Vcx, Vct, Vcxt, Vct.L, Vnw.L, Vscc.L, Vcxt.L)
R> names(vcovs) <- c("OLS", "Vw", "Vcx", "Vct", "Vcxt", "Vct.L", "Vnw.L",
+    "Vscc.L", "Vcxt.L")
```

in order to calculate a comprehensive table of p values from robust estimators:

```
R> cfrtab <- matrix(nrow = length(coef(plmmod)), ncol = 1 + length(vcovs))
R> dimnames(cfrtab) <- list(names(coef(plmmod)), c("Coefficient",
+    paste("s.e.", names(vcovs))))
R> cfrtab[, 1] <- coef(plmmod)
R> for (i in 1:length(vcovs)) {
+    cfrtab[, 1 + i] <- coeftest(plmmod, vcov = vcovs[[i]])[, 2]
+  }
R> print(t(round(cfrtab, 4)))
```

[11] We are grateful to an anonymous reviewer for this observation.

```
            (Intercept) log(pcap) log(pc) log(emp)   unemp
Coefficient      1.6433    0.1550  0.3092   0.5939 -0.0067
s.e. OLS         0.0576    0.0172  0.0103   0.0137  0.0014
s.e. Vw          0.0708    0.0185  0.0125   0.0195  0.0013
s.e. Vcx         0.2442    0.0601  0.0462   0.0686  0.0031
s.e. Vct         0.0944    0.0232  0.0063   0.0246  0.0018
s.e. Vcxt        0.2520    0.0617  0.0450   0.0702  0.0033
s.e. Vct.L       0.1875    0.0461  0.0079   0.0480  0.0031
s.e. Vnw.L       0.1144    0.0299  0.0206   0.0316  0.0020
s.e. Vscc.L      0.1503    0.0370  0.0076   0.0387  0.0025
s.e. Vcxt.L      0.2722    0.0657  0.0389   0.0736  0.0036
```

6.1. PPP regression

This example applies the new combination `Vcxt.L`, which as observed is undocumented in the literature, in the appropriate context of a "long" panel. Its main purpose is to show how to apply the methodology discussed in the paper to linear hypothesis testing.

Coakley, Fuertes, and Smith (2006) present a purchasing power parity (PPP) regression on quarterly data 1973Q1 to 1998Q4 for 17 developed countries, so that $N = 17$ and $T = 104$ which is fairly typical of a "long" panel.[12] The estimated model is

$$\Delta s_{it} = \alpha + \beta(\Delta p - \Delta p^*)_{it} + \nu_{it},$$

where s_{it} is the relative exchange rate against USD and $(\Delta p - \Delta p^*)_{it}$ is the inflation differential between the country and the US.

```
R> data("Parity", package = "plm")
R> fm <- ls ~ ld
R> pppmod <- plm(fm, data = Parity, effect = "twoways")
```

The hypothesis of interest is $\beta = 1$, therefore instead of significance diagnostics we report the corresponding robust Wald test from `linearHypothesis` in package **car** (Fox and Weisberg 2011):

```
R> library("car")
R> linearHypothesis(pppmod, "ld = 1", vcov = Vcxt.L)

Linear hypothesis test

Hypothesis:
ld = 1

Model 1: restricted model
Model 2: ls ~ ld
```

[12] The first of three examples in the original SCC paper (Driscoll and Kraay 1998) is also a purchasing power parity regression, on annual data 1973–1993 for a sample of 107 countries.

```
Note: Coefficient covariance matrix supplied.

  Res.Df Df  Chisq Pr(>Chisq)
1   1648
2   1647  1 2.2942     0.1299
```

6.2. Petersen's artificial data

The last example draws on a well-known simulated dataset, replicating the original results.

To complement his paper, Petersen (2009) produced a simple artificial dataset, which has become an informal benchmark for practitioners. The data can be retrieved from http://www.kellogg.northwestern.edu/faculty/petersen/htm/papers/se/test_data.txt; a copy is provided in the accompanying materials to this article. He provides the following estimates of standard errors: classical, White heteroskedastic, clustered by firm or year, double-clustered by firm and year; and coefficients and standard errors estimated according to the Fama-MacBeth procedure. In the following, I replicate his results in R with **plm**.

```
R> petersen <- read.table(file = "test_data.txt")
R> colnames(petersen) <- c("firmid", "year", "x", "y")
R> ptrmod <- plm(y ~ x, data = petersen, index = c("firmid", "year"),
+    model = "pooling")
R> vcovs <- c(vcov, Vw, Vcx, Vct, Vcxt)
R> names(vcovs) <- c("OLS", "Vw", "Vcx", "Vct", "Vcxt")
R> cfrtab <- matrix(nrow = length(coef(ptrmod)), ncol = 1 + length(vcovs))
R> dimnames(cfrtab) <- list(names(coef(ptrmod)), c("Coefficient",
+    paste("s.e.", names(vcovs))))
R> cfrtab[, 1] <- coef(ptrmod)
R> for(i in 1:length(vcovs)) {
+    cfrtab[, 1 + i] <- coeftest(ptrmod, vcov = vcovs[[i]])[, 2]
+  }
R> print(t(round(cfrtab, 4)))

            (Intercept)      x
Coefficient      0.0297 1.0348
s.e. OLS         0.0284 0.0286
s.e. Vw          0.0284 0.0284
s.e. Vcx         0.0669 0.0505
s.e. Vct         0.0222 0.0317
s.e. Vcxt        0.0646 0.0525
```

One should notice a small difference w.r.t. the results of Petersen: in fact, to replicate them exactly one shall specify to use the same small sample correction Stata (StataCorp. 2015) uses by default: e.g., in the double-clustering case,

```
R> coeftest(ptrmod, vcov = function(x) vcovDC(x, type = "sss"))[, 2]
```

```
(Intercept)           x
 0.06506392  0.05355802
```

which yields the same results as double-clustering in Petersen's example.[13]

7. Conclusions

I have reviewed the different robust estimators for the standard errors of panel models used in applied econometric practice, representing them as combinations of atomic building blocks, which can be thought of as the computational counterparts of statistical objects. In turn, these have been defined, according to the functional orientation of R, as variations of the same general element obtained by choosing a clustering dimension (group or time), a lag order and a function of the residuals (either the element-by-element or the outer product).

While it is feasible to combine these constituents *ad hoc* at user level, the standard estimators used in applied practice (White, Arellano, Newey-West, Driscoll and Kraay SCC, double clustering) are provided under form of predefined combinations ("wrapper" functions) for the sake of user-friendliness. Nevertheless, the user enjoys the freedom to combine elements at will, possibly experimenting with non-standard solutions.

The software framework described is integrated in the R package **plm**, so that composite covariance methods can be applied to objects representing panel models of different kinds (FE, RE, FD and, obviously, OLS). The estimate of the parameters' covariance thus obtained can in turn be plugged into diagnostic testing functions, producing either significance tables or hypothesis tests. A function is a regular object type in R, hence compact comparisons of standard errors from different (statistical) methods can be produced simply by looping on covariance types, as shown in the examples.

An extension to multiple clustering dimensions as in Cameron *et al.* (2011) is ill-suited to bidimensional econometric panels, and hence out of the scope of this paper; it has recently been implemented in package **multiwayvcov** for linear models (Graham *et al.* 2016), and can foreseeably be adapted to panel settings by combining the latter with demeaning functionality in **plm** (i.e., treating the transformed data as a classical linear model, see Section 4.4) in ways that look fairly straightforward but are out of the scope of the present work.

Lastly, one caveat applies. This paper is concerned with design-efficient computing of a quite general class of estimators. Generality will mean that many different estimators can be fitted to the data obtaining numerical estimates. Advances in computing power have made most of these computationally very cheap, hence a conservative "fit-them-all" strategy is feasible (although conservativeness will come at the expense of test power: see the observations at the beginning of Section 6). It must nevertheless be borne in mind that computability does not by any means guarantee statistical soundness, and that the hypotheses under which a covariance estimator is consistent and has desirable properties in finite samples are usually a subset of those under which it is actually computable.

[13] Petersen also reports Fama-MacBeth estimates. As observed in Section 1, these are nothing else but a mean groups estimator where means are taken over time instead of, as is customary in panel time series econometrics, over individual observations. Therefore this last part of Petersen's results can be replicated by swapping indices in the **plm** function `pmg`, as in the following statement: `coeftest(pmg(y ~ x, data = petersen, index = c("year", "firmid")))`.

Acknowledgments

My work has benefited greatly, at different stages, from discussions with: Liviu Andronic; Yves Croissant and Kevin Tappe, my coauthors in **plm**; Achim Zeileis, who originally devised the object-oriented framework for robust inference (Zeileis 2004, 2006); and Mahmood Arai, who provided the original implementation of panel (double-)clustered standard errors in R (Arai 2009). Despite many differences in the software approach, this implementation has been inspired by his work. Useful comments and suggestions from John Driscoll, Harry Kelejian and Douglas Miller are also gratefully acknowledged. Prof. Ron Smith is thanked for the `Parity` data.

References

Anselin L (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publisher, Dordrecht. doi:10.1007/978-94-015-7799-1.

Aptech Systems, Inc (2006). *GAUSS Mathematical and Statistical System 8.0*. Aptech Systems, Inc., Black Diamond, Washington. URL http://www.Aptech.com/.

Arai M (2009). "Cluster-Robust Standard Errors Using R." *Technical report*, Department of Economics and SULCIS, Stockholm University. URL http://people.su.se/ma/clustering.pdf.

Arellano M (1987). "Computing Robust Standard Errors for within Group Estimators." *Oxford Bulletin of Economics and Statistics*, **49**(4), 431–434. doi:10.1111/j.1468-0084.1987.mp49004006.x.

Arellano M (2003). *Panel Data Econometrics*. Oxford University Press.

Bailey D, Katz JN (2011). "Implementing Panel-Corrected Standard Errors in R: The **pcse** Package." *Journal of Statistical Software*, **42**(1), 1–11. doi:10.18637/jss.v042.c01.

Beck N, Katz JN (1995). "What to Do (and Not to Do) with Time-Series Cross-Section Data." *The American Political Science Review*, **89**(3), 634–647. doi:10.2307/2082979.

Beck N, Katz JN (1996). "Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models." *Political Analysis*, **6**(1), 1–36. doi:10.1093/pan/6.1.1.

Berger S, Graham N, Zeileis A (2017). "Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R." *Working Paper 2017-12*, Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, Universität Innsbruck. URL http://EconPapers.RePEc.org/RePEc:inn:wpaper:2017-12.

Cameron AC, Gelbach JB, Miller DL (2011). "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics*, **29**(2), 238–249. doi:10.1198/jbes.2010.07136.

Chudik A, Pesaran MH, Tosetti E (2011). "Weak and Strong Cross-Section Dependence and Estimation of Large Panels." *The Econometrics Journal*, **14**(1), c45–c90. doi:10.1111/j.1368-423x.2010.00330.x.

Coakley J, Fuertes AM, Smith R (2006). "Unobserved Heterogeneity in Panel Time Series Models." *Computational Statistics & Data Analysis*, **50**(9), 2361–2380. doi:10.1016/j.csda.2004.12.015.

Croissant Y, Millo G (2008). "Panel Data Econometrics in R: The **plm** Package." *Journal of Statistical Software*, **27**(2), 1–43. doi:10.18637/jss.v027.i02.

Driscoll JC, Kraay AC (1998). "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *The Review of Economics and Statistics*, **80**(4), 549–560. doi:10.1162/003465398557825.

Fama EF, MacBeth JD (1973). "Risk, Return, and Equilibrium: Empirical Tests." *The Journal of Political Economy*, **81**(3), 607–636. doi:10.1086/260061.

Fox J, Weisberg S (2011). *An R Companion to Applied Regression*. 2nd edition. Sage, Thousand Oaks.

Franzese R (1996). "A GAUSS Procedure to Estimate Panel-Corrected Standard-Errors with Non-Rectangular and/or Missing Data." *The Political Methodologist*, **7**(2), 2–3.

Froot KA (1989). "Consistent Covariance Matrix Estimation with Cross–Sectional Dependence and Heteroskedasticity in Financial Data." *Journal of Financial and Quantitative Analysis*, **24**(3), 333–355. doi:10.2307/2330815.

Golden MA, Picci L (2008). "Pork-Barrel Politics in Postwar Italy, 1953–94." *American Journal of Political Science*, **52**(2), 268–289. doi:10.1111/j.1540-5907.2007.00312.x.

Graham N, Arai M, Hagströmer B (2016). "**multiwayvcov**: Multi-Way Standard Error Clustering." R package version 1.2.3, URL https://CRAN.R-project.org/package=multiwayvcov.

Greene WH (2003). *Econometric Analysis*. 5th edition. Prentice Hall.

Hothorn T, Hornik K, Van De Wiel MA, Zeileis A (2006). "A Lego System for Conditional Inference." *The American Statistician*, **60**(3), 257–263. doi:10.1198/000313006x118430.

Ibragimov R, Müller UK (2010). "*t*-Statistic Based Correlation and Heterogeneity Robust Inference." *Journal of Business & Economic Statistics*, **28**(4), 453–468. doi:10.1198/jbes.2009.08046.

Liang KY, Zeger SL (1986). "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika*, **73**(1), 13–22. doi:10.1093/biomet/73.1.13.

MacKinnon JG, White H (1985). "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics*, **29**(3), 305–325. doi:10.1016/0304-4076(85)90158-7.

Munnell A (1990). "Why Has Productivity Growth Declined? Productivity and Public Investment." *New England Economic Review*, **January/February**, 3–22.

Newey WK, West KD (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, **55**(3), 703–708. doi:10.2307/1913610.

Pesaran MH, Smith R (1995). "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics*, **68**(1), 79–113. doi:10.1016/0304-4076(94)01644-f.

Petersen MA (2009). "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *The Review of Financial Studies*, **22**(1), 435–480. doi:10.1093/rfs/hhn053.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

StataCorp (2015). *Stata Data Analysis Statistical Software: Release 14*. StataCorp LP, College Station. URL http://www.stata.com/.

The MathWorks Inc (2017). *MATLAB – The Language of Technical Computing, Version R2017a*. Natick. URL http://www.mathworks.com/products/matlab/.

Thompson SB (2011). "Simple Formulas for Standard Errors That Cluster by Both Firm and Time." *Journal of Financial Economics*, **99**(1), 1–10. doi:10.1016/j.jfineco.2010.08.016.

White H (1980). *Asymptotic Theory for Econometricians*. Academic Press, Orlando.

White H (1984). "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity." *Econometrica*, **48**(4), 817–838. doi:10.2307/1912934.

Wilson SE, Butler DM (2007). "A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications." *Political Analysis*, **15**(2), 101–123. doi:10.1093/pan/mpl012.

Wooldridge JM (2002). *Econometric Analysis of Cross-Section and Panel Data*. MIT Press.

Zeileis A (2004). "Econometric Computing with HC and HAC Covariance Matrix Estimators." *Journal of Statistical Software*, **11**(10), 1–17. doi:10.18637/jss.v011.i10.

Zeileis A (2006). "Object-Oriented Computation of Sandwich Estimators." *Journal of Statistical Software*, **16**(9), 1–16. doi:10.18637/jss.v016.i09.

Zeileis A, Hothorn T (2002). "Diagnostic Checking in Regression Relationships." *R News*, **2**(3), 7–10.

Zeileis A, Lumley T (2017). ***sandwich**: Robust Covariance Matrix Estimators*. R package version 2.4-0, URL https://CRAN.R-project.org/package=sandwich.

Affiliation:

Giovanni Millo
Group Insurance Research, Generali SpA
and
DEAMS, University of Trieste
Via Machiavelli 3
34132 Trieste, Italy
Telephone: +39/040/671184
Fax: +39/040/671160
E-mail: Giovanni.Millo@Generali.com

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
November 2017, Volume 82, Issue 3 *Submitted:* 2015-12-15
doi:10.18637/jss.v082.i03 *Accepted:* 2016-10-12