ECONOMICS OF REGULATION AND ANTITRUST

Fourth Edition

W. Kip Viscusi
Joseph E. Harrington, Jr.
John M. Vernon

The MIT Press
Cambridge, Massachusetts
London, England

enhanced efficiency. They provide financial rewards to achieving improvements in specific operating measures.

Here is an example that has been widely applied to electric utilities since the 1970s[5]:

> These incentive payment programs take many forms and focus on different operating statistics: they reward utilities which experience high levels of base load generating unit utilization and availability, low heat rates (reflecting the efficient transformation of fuel into electricity), and keep fuel and purchased power costs below externally-determined indices.

A firm that has a heat rate below the specified level will be able to retain some or all of the cost savings. In the case of New York, the regulatory agency forecast the rise in fuel costs and allowed the electric utility to pass on to consumers only 60 to 80 percent of any costs above that forecast; in addition, the utility retained a similar fraction of any savings when costs fell below that forecast. This created incentives to keep fuel costs down, incentives that were lacking when there was an automatic fuel cost pass-through program.

Though a regulatory agency may not jettison rate-of-return regulation for the incentive-based methods we are about to describe, this example highlights how incentives can be created in a more focused and less encompassing way alongside traditional regulation. Interestingly, one study found that these more focused methods tended to result in greater gains in efficiency than earnings sharing schemes, which we review next.[6]

Earnings Sharings

The basic problem with rate-of-return regulation is that the firm does not get to share in any of the cost savings; reductions in cost induce commensurate reductions in price so as to keep the firm's rate of return fixed. The regulated firm then has no incentive to lower cost (except that which is created due to regulatory lag).

Earnings sharing (or sliding scale) regulation is based on the idea that if we want to induce regulated firms to reduce cost and engage in other efficiency-enhancing practices, we need to allow them to retain some of the gains that they create. However, if we let them retain all of the gains, then we are, in essence, leaving them unregulated. Earnings sharing regulation finds middle ground by having the firm and consumers share in any excess earnings.

As an example, consider the earnings sharing scheme that Pacific Bell faced in the 1990s in California. It could keep all profits if its rate of return was less than 13 percent. If the rate of return was between 13 percent and 16.5 percent, it could retain 50 percent of the profits in excess of the 13 percent return, with the remainder being rebated to its customers. Finally,

5. Sanford V. Berg and Jinook Jeong, "An Evaluation of Incentive Regulation for Electric Utilities," *Journal of Regulatory Economics* 3 (March 1991): 45–55. The quotation is from page 45.

6. Christopher R. Knittel, "Alternative Regulatory Methods and Firm Efficiency: Stochastic Frontier Evidence from the U.S. Electricity Industry," *Review of Economics and Statistics* 84 (August 2002): 530–40.

all profits in excess of a 16.5 percent return were to be rebated. This resulted in a maximum rate of return of 14.75 percent.

A general class of earnings sharings can be described by the following formula. Letting r denote the gross rate of return (that is, before netting out customers' share of profit), a firm's net rate of return is:

$$\begin{cases} r & \text{if } r \le \underline{r} \\ \underline{r}+\theta(r-\underline{r}) & \text{if } \underline{r} \le r \le \bar{r} \\ \underline{r}+\theta(\bar{r}-\underline{r}) & \text{if } \bar{r} \le r \end{cases} \tag{12.5}$$

where $\underline{r} < \bar{r}$ and $0 \le \theta \le 1$. When its gross rate of return lies in the band ranging from $\underline{r}$ to $\bar{r}$, the regulated firm retains a fraction θ of the excess profit. But its rate of return is capped at $\underline{r} + \theta(\bar{r} - \underline{r})$ as it must rebate all excess profit once a gross rate of return of $\bar{r}$ is reached. In the Pacific Bell example, $\underline{r} = 0.13$, $\bar{r} = 0.165$, and $\theta = 0.5$.

The higher is θ, the greater is the incentive for the firm to reduce cost and raise revenue since it is able to retain a higher fraction of the rise in profit. Of course, it also means that it is able to set higher prices, so one would not want to set θ and $\bar{r}$ too high or the entire reason for regulation would be mooted. Notice that this class of regulatory schemes includes some of the previous cases we have discussed. Traditional rate-of-return regulation is when $\theta = 0$ and $\underline{r}$ is the allowed rate of return, while the case of an unregulated monopoly is when $\theta = 1$ and $\bar{r} = \infty$.

Example: San Diego Gas & Electric

For a concrete example of incentive regulation, consider San Diego Gas & Electric in 1999.[7] The earnings sharing scheme is depicted in figure 12.2 and has ten steps to it, a bit more complicated than what was described above. The authorized rate of return is 9.05 percent and there is a "deadband" of 0.25 percent so that SDG&E keeps all returns up to 9.30 percent. This is not so much an incentive device as a simplification in light of the unpredictability of profits. For a gross rate of return between 9.30 percent and 9.80 percent, SDG&E keeps 25 percent of the additional profit, and the share that it retains goes up with each step to the point that it keeps all additional profits once the gross rate of return is 12.05 percent or higher. This scheme then provides increasingly high-powered incentives as profits rise.

In addition to earnings sharings, the regulatory body provided rewards and penalties concerning performance with respect to employee safety, customer satisfaction, telephone response time, and system reliability that, in sum, could result in an annual penalty or reward of as much as $14.5 million. For example, the benchmark for telephone response time is that

Price

7. This discussion is based on Richard Myers and Laura Lei Strain, "Electric and Gas Utility Performance Based Ratemaking Mechanisms" (Energy Division, California Public Utilities Commission, September 2000).

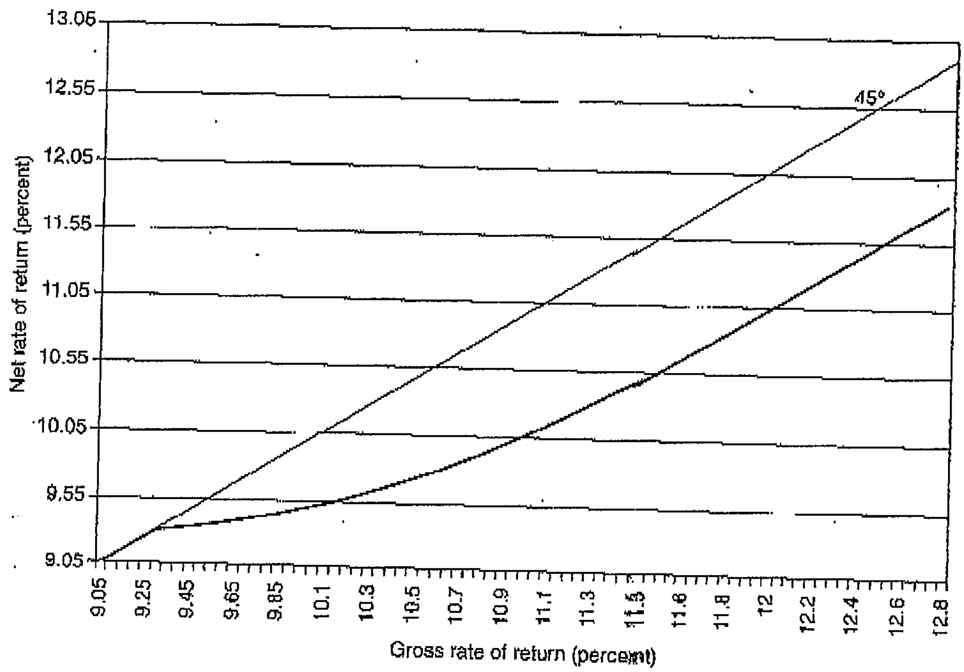

Figure 12.2
Allowed Rate of Return for San Diego Gas & Electric

80 percent of calls are answered in 60 seconds and there is a reward (penalty) of $10,000 for each 0.1 percent increase (decrease) above (below) 80 percent, with a maximum reward or penalty of $1.5 million. Table 12.1 shows the results for 1999.

Price Caps

Price caps were first proposed in the early 1980s by Stephen Littlechild in connection with the regulation of British Telecom and Roy Radner at AT&T's Bell Labs. They are based on the idea that if the price the firm can charge is independent of any cost reductions, then it knows that any such reductions will go to the firm's bottom line. This provides powerful incentives to act efficiently and, in addition, gives the firm some flexibility in adjusting price. As opposed to earnings sharing, the constraint is on price rather than profit. The trick to price caps is setting price at an appropriate level, which requires forecasting future productivity gains.

A policy of price caps requires that the regulatory agency specify a maximum price, which is adjusted on a predetermined frequency according to a predetermined formula. This formula