Disallowances and Overcapitalization in the U.S. Electric Utility Industry

Stratford Douglas, Thomas A. Garrett, and Russell M. Rhine

Regulation of an industry often produces unintended consequences. Averch and Johnson (1962) argue that certain regulation of electric utilities provides utilities the incentive to purchase an inefficiently large amount of capital. Another possible and related unintended consequence of electric utility regulation is that regulatory cost disallowances on capital may also increase utilities’ incentives to overcapitalize. The authors provide theoretical evidence that capital expenditure disallowances will increase the Averch and Johnson effect in some instances and thus may have contributed to the overcapitalization problem that regulation was designed to discourage. Our model shows that disallowances can reduce the rate of return on investment and thereby increase the Averch and Johnson distortion. (JEL D42, L43, L51)

The industry increased output by 20 percent from 1995 to 2006, and generation capacity is expected to grow by another 8 percent over the next five years. Currently, nearly 98 percent of the existing capacity consists of fossil fuel power plants, nuclear reactors, hydroelectric power plants, and other renewable energy sources. Although these sources all contribute to the total generation, fossil fuels generate the majority of electricity. Natural gas, coal, and petroleum supply 41 percent, 31 percent, and 6 percent of generation capabilities, respectively. Nuclear, hydroelectric, and other renewable sources comprise approximately 19 percent. To date, the academic literature has devoted much attention to the U.S. electric

All households, firms, and government entities depend on one or more of the 3,170 electric utilities in the United States to provide a reliable source of energy. These electric companies consist of investor-owned, publicly owned, cooperative, and federal utilities. Only 8 percent of these utilities are investor owned, but they produce approximately 75 percent of the total generating capability. Publicly owned and federal utilities each generate about 10 percent of the country’s electricity, and cooperatives generate the remaining 4 percent. The revenue from retail sales (to ultimate consumers) for all electric utilities amounted to $326 billion in 2006 and represents about 2.5 percent of gross domestic product.

1 Investor-owned utilities are private corporations that operate to produce a rate of return for their investors. Publicly owned utilities are nonprofit agencies owned by local governments. Cooperatives are owned by members of a community and typically operate in rural areas where investor-owned utilities are not economically feasible. Federal electric utilities are owned and operated by the federal government.

2 Other renewable sources, as defined by the Energy Information Administration (2007), include wood, black liquor, other wood waste, municipal solid waste, landfill gas, sludge waste, tires, agriculture by-products, other biomass, geothermal, solar thermal, photovoltaic energy, and wind.

3 For details, see Energy Information Administration (2007).
utility industry. The primary reason for such interest is that electricity is used by all Americans, and firms in the industry enjoy a monopolistic market structure, at least at the distribution level. Although the academic literature is broad in scope, most articles fall into two categories. The first category is cost analysis—primarily the measuring of scale economies. That is, researchers attempt to determine where firms are operating on their long-run average cost curves and subsequently determine whether production costs can be lowered by having firms increase or decrease their scale of production. The second category, much larger than the first, is analysis of the regulatory aspect of the industry and the unanticipated consequences of those regulations. Relevant regulations involve not only those related to the environmental impact of electricity generation but also those regulating profits by setting the price that firms are allowed to charge for their electricity.

One specific issue that has sparked much attention is the overcapitalization of the electric utility industry—that is, electric utilities hold a quantity of capital that is greater than the cost-minimizing quantity. Averch and Johnson (1962; hereafter A-J) argued that privately owned utilities invest in capital beyond the cost-minimizing level in response to the incentives offered by regulation. The authors showed how a regulator, by tying a firm’s allowed profit to its capital stock and offering a rate of return on capital that exceeds the marginal cost of capital, provides the firm the incentive to purchase an inefficiently large amount of capital. The A-J model has been thoroughly analyzed, discussed, and tested in the academic literature.

The key aspect of the A-J model is that regulation of electric utilities results in the unintended consequence of overcapitalization. We argue in this paper that the A-J model is not complete and can thus be expanded to account for another aspect of electric utility regulation: A regulator’s denial of cost recovery for some portion of the utility’s capital expenditures provides an additional incentive for firms to overcapitalize. The basis for this incentive is that a portion of the firm’s capital is excluded from the calculation of profit by the regulator. Firms will thus invest more in capital to maintain or increase profits.

One purpose of the regulatory disallowances was to make utility management accountable for cost overruns and thereby reduce their incentive to overcapitalize. Lyon and Mayo (2005) state that the disallowances were punitive and directed toward poorly managed firms. The disallowances succeeded in the sense that they apparently reduced utilities’ appetite for constructing large new power plants; few large power plants and no new nuclear plants have been initiated in the past 20 years. But a subtler question remains as to whether the disallowances increased the efficiency of utilities’ capital purchase decisions given the higher cost of capital.

We preface our theoretical framework by first providing an overview of the economic effects of regulation, including unintended consequences. We then provide a conceptual framework for overcapitalization that serves as a basis for our theoretical model, which expands the A-J model. We find theoretical support for the proposition that regulatory cost disallowances increase utilities’ incentives to overcapitalize. Thus, the overcapitalization in the electric utility industry is a result not only of the A-J effect, but also of regulatory denial of cost recovery for a portion of a utility’s capital expenditures.
As demonstrated in Figure 1, natural monopolies, by definition, exhibit decreasing average and marginal costs over a wide range of output because of high fixed costs (plants, equipment) and low variable and marginal costs. Thus, one firm can produce most or all of the electricity demanded by consumers more cheaply than could multiple firms. Monopoly pricing involves charging a price \( P_M \) greater than the marginal cost \( MC \) of production and producing an output level \( Q_M \) less than that under perfect competition \( QC \), therefore resulting in a loss in economic efficiency.

The basic model of monopoly regulation posits that regulators aim to reduce the price charged by the monopolist and expand the monopolist’s output. One common approach is to set the price equal to the marginal cost of production (point \( E \)). This mimics pricing under perfect competition. However, given the cost structure of monopoly, marginal costs are below average costs \( AC \), so marginal cost pricing often results in a financial loss for the monopolist. Average cost pricing deviates from the competitive price and output level (because average costs > marginal costs), but still results in a price and output level (point \( F \)) that approximates the competitive solution.

Although in theory the regulation of monopoly pricing is fairly straightforward, in reality it may be difficult to achieve the price and output levels that would exist under perfect competition, given that regulation occurs in political markets. As first discussed by Stigler (1971), consumers and producers have different objectives with regard to monopoly prices—consumers prefer lower prices and greater output, whereas the monopolist prefers higher prices and lower output. According to Stigler, it is reasonable to assume that both groups exert political pressure to set regulatory outcomes in their favor.

Stigler’s model shows that a vote-maximizing regulator will set a utility’s sale price of electricity such that the marginal gain in support from producers is just offset by the loss in consumer votes. Thus, in Figure 1, political competition between consumers and producers will ensure that the regulated price will lie somewhere between \( P_M \) and \( P_C \) and the regulated output will lie somewhere between \( Q_M \) and \( Q_C \). The exact location will depend on the relative strengths of consumer and producer groups in exerting political pressure. Given that the cost of organizing producer groups is much less than the cost of organizing thousands or millions of consumers (Olson, 1965; Peltzman, 1976), producers are likely to exert more political pressure than are consumers; as a result, regulation will likely favor producers.

Another aspect of regulation that has garnered attention in the literature—and is most relevant for the purpose of our paper—is the potential for unintended consequences as a result of regulation. Unintended consequences are, as the term suggests, unanticipated effects from policy actions. Examples from other industries include those described by Hall, Propper, and Reenan (2008), who find that regulated pay for medical staff across geographically heterogeneous labor markets results in problems with recruiting, retaining, and motivating high-quality workers, which ultimately

---

5 See Warkentin-Glenn (2006) for a history of the electric utility industry.

6 The cost curves shown in Figure 1 are for a representative natural monopolist and may not represent the identical cost and pricing structure of all monopolists.
affects hospital performance; and Nelson (2003), who finds that the regulation of alcohol advertising initiated to restrict demand for one type of product creates increased demand for other alcoholic products.

As mentioned earlier, overcapitalization in the electric utility industry is one unintended consequence of regulating the industry. Specifically, a firm is motivated to purchase an inefficiently large amount of capital because a regulator ties the firm’s allowed profit to its capital stock. We argue that overcapitalization will be exacerbated when a portion of a firm’s capital is excluded from the calculation of profit by the regulator, and firms will thus invest more in the type of capital not excluded from the profit calculation to maintain or increase profits.

Overcapitalization Explained

Many industry observers, both inside and outside academia, believe that the electric utility industry in the United States is overcapitalized. Our theoretical model presented in the next section focuses on overcapitalization, and so we discuss overcapitalization in general to better lay the foundation for our theoretical model.

In Figure 2, $K^*$ is the cost-minimizing capital stock for production along the ex post isoquant. Suppose we observe a firm producing inefficiently at point $B$ using capital stock, $K$. The A-J model explains a firm’s decision to produce at point $B$ as a rational response to regulatory incentives. Once it arrives at $B$, it will tend to stay there. More intuitively, the price that an electric utility is allowed to charge for its output is set by industry regulators. The regulators choose a price so that the firm will generate enough revenue to cover the operating costs, which include labor expenses, and to generate enough profit to pay the investors a fair rate of return on capital. Thus, with production inputs of capital and labor, the firm knows that an increase in labor expenses will be offset by an increase in revenues by the same amount. However, an increase in capital investment will be offset by an increase in revenue that is greater than the increased costs to ensure that investors receive a fair rate of return. Thus, the firm has an incentive to increase its investment in capital. Our theoretical exercise in the next section of the paper determines whether a firm will choose point $B$ instead of $A$ because of regulatory incentives or for some other reason.

Overcapitalization can appear in many different guises. Utilities that need new capacity might avoid leasing available power plants, preferring to build their own. They might resist selling power plants that they do not need. They might resist joining efficient power pools. They might inefficiently choose capital-intensive expenditures for pollution abatement (such as scrubbers) over less capital-intensive alternatives (such as co-firing or “green power” purchases). They also might appear overeager to move to underground distribution systems. They might impose excessive safety and reliability standards on themselves, resulting in a “reserve margin” in the form of an idle generation plant that exceeds any rational requirements. As Kahn (1988) explains, utilities may have “excessively high (because excessively costly) standards of reliability and uninterruptibility of service, with correspondingly high and costly specification for the equipment they employ.” Finally, utilities may allow themselves to be overcharged for capital equipment, as occurred in the
electrical equipment manufacturers’ price conspiracy of the 1950s (Emery, 1973).

Overcapitalization also can occur because of forecast error when the firm overestimates future demand. Faulty forecasting is a reasonable suspect as a cause of utility overcapitalization. The late 1970s and early 1980s saw both a decrease in the U.S. rate of economic growth and, because of the increase in relative energy prices, the first significant divergence between rates of growth in the economy and in the demand for energy. Traditional methods of demand forecasting failed to account for these trends, and utilities continued to build large power plants to serve demand that, in many cases, never materialized. Exacerbating this trend was the greatly increased cost of nuclear power as a result of the regulatory response to the Three Mile Island incident in 1979.

Large generation plants must be built well before anticipated demand is realized. If the ex ante isoquant in Figure 2 were the anticipated level of demand, then \( K \) would be the ex ante cost-minimizing level of capital. However, when actual demand is realized ex post, the firm finds itself operating off its least-cost expansion path at point \( B \). If presented with efficient incentives, the firm would move to point \( A \) as quickly as possible. Thus, if overcapitalization arose solely in response to overestimation of demand, then overcapitalization should decline over time as the firm adapts to the unexpectedly low-demand environment.

Past studies on adapting to a world in which regulators occasionally deny recovery or return on part or all of a utility’s capital investment have shown varying results. Lyon (1991), Gal-Or and Spiro (1992), and Gilbert and Newbery (1994) argue that capital disallowances may decrease investment, whereas Tiesberg (1993) finds that capital disallowances may increase investment. The disallowance might be expected to increase the riskiness of the firm’s investments, and this increased riskiness of capital could cause utilities to reduce their use of capital. If so, the model presented in this paper identifies a countervailing incentive—a tendency to overcapitalize even further because of the disallowance effect on a utility’s de facto allowed rate of return.

### The A-J Model

The A-J model depicts a profit-maximizing firm that is subject to a regulated rate of return. The production function specification may be inappropriate for the traditional electric utility, given the generally accepted stylized facts of the industry. A production function assumes that a firm maximizes output for a given level of inputs; however, a traditional electric utility is more likely to find its output level fixed because its price is regulated within its geographically defined service area. It will choose its input quantities endogenously in response to largely exogenous factor market prices. Therefore, the dual cost function, in which input quantities are chosen in response to fixed factor prices and output levels, is a better model for the behavior of the electric utility.

A-J demonstrate that if the allowed rate of return exceeds the cost of capital, then a firm’s capital stock will increase to the point where the cost of capital exceeds its marginal product. Let \( R(Y), P, Y, P_i, K, L, \) and \( s_K \) be the revenue function, output price, the quantity of output, the price of the inputs, the quantity of capital, the quantity of labor, and the allowed rate of return, respectively, where \( i = K, L \). The price of capital, \( P_K \), is the interest cost in holding plant and equipment, which differs from the acquisition cost of capital, \( c \). We assume that \( c = 1 \); thus, the regulator constrains the firm to an allowed rate of return, \( s_K \), that will cover the physical quantity of capital, \( c \cdot K \), after operating expenses are deducted from revenue. That is,

\[
\frac{PY - PL}{K} \leq s_K.
\]

Equation (1) below is the Lagrangian for a profit-maximizing firm. The firm maximizes profits subject to two constraints. The first constraint is the regulatory-constrained allowed rate of return, and the second constraint is output that is subject to the production function, \( Y \leq F(K, L) \). Equations (2) through (6) are the first-order conditions that result from maximizing equation (1):

\[
\mathcal{L}(K, L, \lambda, \mu) = R(Y) - P_K K - P_L L + \lambda(s_K K - R(Y) + P_L L) + \mu(F(L, K) - Y)
\]
We assume that $R' > 0$, $F_L > 0$, $F_K > 0$ and that the firm earns a profit (i.e., $s_K > P_K$). We also set $G = R(F(K,L))$ and assume that it is concave so that the second-order condition for maximization is satisfied. Therefore, the marginal revenue products of labor and capital are $G_L = R'F_L$ and $G_K = R'F_K$, respectively. Finally, we assert that $1 > \lambda > 0$ and $\mu > 0$ (Takayama, 1993). By combining equations (2) and (5), it follows that the firm uses labor efficiently: $G_L$ equals the wage; but by combining equations (4) and (5), it follows that firms do not use capital efficiently: $G_K$ is less than $P_K$, as shown in equation (7):

$$G_L = P_L, \quad G_K < P_K.$$

By totally differentiating equation (6) with respect to $s_K$ and applying the efficient use of labor from equation (7), the A-J effect can be produced and is shown in equation (8).\(^7\) Decreasing the allowed rate of return results in the firm increasing its capital stock:

$$K + s_K \frac{dK}{ds_K} - G_K \frac{dK}{ds_K} + P_L \frac{dL}{ds_K} - G_L \frac{dL}{ds_K} = 0 \Rightarrow K = (G_K - s_K) \frac{dK}{ds_K} + (G_L - P_L) \frac{dL}{ds_K},$$

$$\Rightarrow \frac{dK}{ds_K} = \frac{K}{G_K - s_K} < 0.$$

The capital stock, $K$, plays a twofold role in a regulated firm, as both a productive input and the rate base for computing allowed profit. The rate base is defined as the amount of capital expenditures that the regulator uses to determine the costs that should be passed on to consumers. The regulator determines whether expenditures on capital may be included in the firm’s rate base and therefore paid for by utility customers. If a portion of a firm’s capital expenditure was not prudently incurred, then the regulator may disallow that portion (i.e., exclude it from the rate base). Many utilities incurred massive disallowances in the 1980s, primarily because of cost overruns and tighter safety standards for nuclear power plants.

We suggest that the disallowances in fact exacerbated the A-J effect in some instances, thereby reducing the efficiency of capital use. The exact effect depends on whether the disallowed power plant is capable of producing electricity for sale. Suppose that the regulator disallows a capital expenditure on a power plant that represents some proportion, $\delta$, of the utility’s total rate base, where $0 < \delta < 1$. If the power plant is never completed or is not allowed to operate, then the full capital expenditure appears as a cost, but only the allowed portion appears in the rate base and the production function.\(^8\) Thus, the disallowed capital expenditure is useless—it is not generating electricity or revenues for the firm. The Lagrangian for disallowed useless capital is shown by equation (9):

$$\mathcal{L}_{(\text{useless})}(K, L, \lambda, \mu) = R(Y) - P_K K - P_L L + \lambda(s_K (1 - \delta) K - R(Y)) + P_L L + \mu(F((1 - \delta) K, L) - Y).$$

where $R(Y) \equiv G((1 - \delta) K, L)$.

The firm incurs the full cost of purchasing $K$, but the allowed rate of return, $s_K$, applies only to the

---

\(^7\) See Takayama (1993, pp. 215-16) for a more detailed discussion and rigorous proof.

\(^8\) An example of a completed power plant that was unable to be used for production is the Shoreham nuclear power plant in New York; on completion, it was never allowed to operate because of safety concerns.
rate base \((1 - \delta)K\) and only \((1 - \delta)K\) is productive. Essentially, the disallowance reduces the firm’s capital stock, both as productive input and as a portion of the rate base, while retaining a sunk cost of \(P_k \delta K\) on the firm. Equations (10) through (14) are the first-order conditions that result from the maximization of equation (9):

\[
(10) \quad \frac{\partial L}{\partial L} = -P_L (\lambda - 1) + \mu F_L = 0
\]

\[
(11) \quad \frac{\partial L}{\partial \mu} = F((1 - \delta)K, L) - Y = 0
\]

\[
(12) \quad \frac{\partial L}{\partial K} = -P_K + \lambda (1 - \delta)s_K + (1 - \delta)\mu F_K = 0
\]

\[
(13) \quad \frac{\partial L}{\partial Y} = (1 - \lambda)R' - \mu = 0
\]

\[
(14) \quad \frac{\partial L}{\partial \lambda} = s_K (1 - \delta)K - R + P_L L = 0.
\]

The same first- and second-order conditions as previously stated hold, and the firm earns a profit, \((1 - \delta)s_K > P_K\). Combining equations (10) and (13) indicates efficient use of labor, as shown in equation (15):

\[
(15) \quad G_L = P_L.
\]

The incremental input distortion caused by a change in the allowed rate of return, \(s_K\), is the same in the model with a useless capital disallowance as it is in the classic A-J model. This can be seen by totally differentiating equation (14) with respect to \(s_K\) and applying equation (15). Equation (16) shows the effect of a change in \(s_K\) on the capital stock, \(K\):

\[
(16) \quad \frac{\delta K}{\delta s_K} = \frac{\frac{dK}{ds_K} + (1 - \delta)s_K \frac{dL}{ds_K} - (1 - \delta)G_K \frac{dK}{ds_K}}{s_K} = \frac{P_L}{s_K} \frac{dL}{ds_K} - G_L \frac{dL}{ds_K} = 0
\]

\[
\Rightarrow (1 - \delta)K = (G_K - s_K)(1 - \delta)\frac{dK}{ds_K} + (G_L - P_k) \frac{dL}{ds_K}
\]

\[
\Rightarrow \frac{dK}{ds_K} = \frac{K}{G_K - s_K} < 0.
\]

Although no change occurs in the marginal effect of an increase in the allowed rate of return, an increase in the disallowance of useless capital does, in fact, increase the amount of capital purchased. Equation (17) is the result of totally differentiating equation (14) with respect to \(\delta\), and it indicates that an increase in the scale of the disallowance causes an increase in \(K\):

\[
(17) \quad \frac{dK}{d\delta} = \frac{s_KK}{(1 - \delta)(s_K - G_K)} > 0.
\]

More commonly, the disallowed capital is capable of producing electricity (i.e., it is “useful”), but it is either not needed to serve the utility’s captive customer base or its construction costs are judged excessive. In the latter case, the utility will take the disallowed costs out of its rate base, but it may continue to operate and sell the power either to its customers or off-system on the wholesale market. Thus, as shown in equation (18), the disallowed unit remains in the production function

\[
\begin{align*}
&\mathcal{L}_{\text{useful}}(K, L, \lambda, \mu) = R(Y) - P_K - P_L L \\
&\quad + \lambda (s_K (1 - \delta)K - R + P_L L) \\
&\quad + \mu (F(K, L) - Y),
\end{align*}
\]

where \(R(Y) = G(K, L)\).

Because the first-order conditions with respect to \(Y\) and \(L\) are identical to that of \(\mathcal{L}_{\text{useless}}, G_L = P_L\) still holds. However, the remaining first-order conditions do differ and are shown as equations (11’), (12’), and (14’):

\[
(11') \quad \frac{\delta \mathcal{L}}{\delta \mu} = F(K, L) - Y = 0
\]

\[
(12') \quad \frac{\delta \mathcal{L}}{\delta K} = -P_K + \lambda (1 - \delta)s_K + \mu F_k = 0
\]

\[
(14') \quad \frac{\delta \mathcal{L}}{\delta \lambda} = s_K (1 - \delta)K - R + P_L L = 0.
\]

The effect of changing the allowed rate of return, \(s_K\), is somewhat different in the presence of useful but disallowed capital,
The inequality in equation (19) holds if the post-disallowance rate of return exceeds the marginal revenue product of capital: that is, 

\[
\frac{(1 - \delta)K}{G_K - (1 - \delta)s_K} < 0.
\]

The marginal distortive effect of increasing \(\delta\) for useful capital (equation (21)), exceeds that of useless capital (equation (17)), by the same factor as in equation (20); that is, \([(1 - \delta)(s_K - G_K)]/[1 - \delta s_K - G_K]\).

Thus, the modified A-J model predicts that a firm subject to a capital disallowance will overcapitalize to a greater degree than a firm without a disallowance. More intuitively, our results indicate that because a utility’s profit is equal to \([(1 - \delta)P_AR - P_K]K\), the existence of disallowed capital, \(\delta\), will cause the utility’s profit to decrease, all other things equal. For the firm to regain those lost profits, either its capital stock, \(K\), or the allowed rate of return, \(s_K\), must increase. Consequently, the only way to regain the profits is by investing in capital because its allowed rate of return cannot be changed by the firm, as it is set by the regulators.

In addition, a firm with disallowed capital that is operable will experience a greater A-J distortion than a firm with disallowed capital that is not operable or a firm with no disallowed capital. This difference is due to the fact that the operable disallowed capital continues to generate revenue for the firm; subsequently, the marginal revenue product of capital is not reduced by the existence of capital that generates no revenue.

**CONCLUSION**

We provide theoretical evidence that capital expenditure disallowances reinforce the Averch-Johnson effect and thereby may have contributed to the overcapitalization problem they were designed to decrease. The theoretical model shows that disallowances, especially of useful plants, reduce the de facto allowed rate of return on capital and thereby increase the Averch-Johnson input distortion. All these results support the idea that overcapitalization is a profit-maximizing response to rate-of-return regulatory incentives and not a mistake caused by incorrect demand forecasting. These results have some relevance to current policy. Regulators and regulated firms in several industries, including natural gas, electricity, and telecommunications, are taking capital assets out of their rate bases. Currently, the reason for such action is not imprudence but deregulation. Our results suggest that the way in which this process occurs matters. In particular, taking the assets out of the rate base, but leaving them in the hands of a regulated firm that is subject to an overall rate-of-return constraint, could result in inefficient overcapitalization. In general, regulators need to be aware that their policies could have unanticipated consequences.

**REFERENCES**


