Chapter IV.5

THE EFFECTIVENESS OF RATE-OF-RETURN REGULATION: AN EMPIRICAL TEST USING PROFIT FUNCTIONS*

THOMAS G. COWING

State University of New York

1. Introduction

In the last decade considerable progress has been made in investigating both the theoretical properties and extensions of the profit-maximizing firm subject to rate-of-return regulation in the form of a constraint on its earned rate of return, the so-called Averch–Johnson model. More recently, several econometric studies of regulatory effectiveness within the general framework of the A–J model have appeared.[1] While much of the theoretical work has been useful in terms of increasing our understanding of the A–J effect, the econometric studies to date have generally suffered from two major defects. The first is a failure to develop econometric models which are fully consistent with the implications of the theoretical model. The second is a failure to adequately treat the problem of simultaneous equation bias.

The purpose of this chapter is to use the theory of profit functions to

*This paper was written during the author's tenure as a Brookings Economic Policy Fellow at the Federal Energy Administration. He would like to especially thank David Nissen of the FEA for his generous encouragement and support. He would also like to acknowledge helpful comments from M. Fuss, D. McFadden, D. Nissen, V. Kerry Smith and members of the Applied Microeconomic Theory Workshop at SUNY-Binghamton. The views and conclusions expressed in this study are those of the author, as is the responsibility for remaining errors, and should not be attributed to any of the agencies or institutions mentioned above.

[1]See Spann (1974), Courville (1974), and Petersen (1975).

derive an econometric framework which is both theoretically compatible with the static Averch–Johnson model and yields consistent estimators of the underlying technological parameters. The next section of this chapter briefly reviews the Averch–Johnson model of the regulated firm, in particular, the comparative statics properties of the model. In Section 3, a generalized profit function for a profit-maximizing regulated firm facing given output and input prices and a maximum allowed rate of return on capital set by a regulatory commission is derived. Although Hotelling's lemma with respect to the partial derivatives of the unconstrained profit function does not hold, a revised version of this lemma is derived for the case of a regulated profit function and is used to develop the econometric model used in this paper. An interesting property of this econometric model is that it allows λ, the Lagrangian multiplier associated with the rate-of-return constraint and, hence, a measure of regulatory effectiveness, to be treated as an endogenous variable and to be estimated even though λ is obviously unobservable. The result is that estimates of λ are allowed to vary across firms, a property which follows directly from the theoretical model but which previous empirical studies of the regulated firm have failed to include.

Several statistical tests of regulatory effectiveness are derived using the regulated profit function, and the results of these tests are presented in the fourth section. The data used in this analysis consist of *ex ante* observations on 114 new steam-electric plants owned by regulated privately-owned U.S. electric utilities, and installed between 1948 and 1965. The results of the first test, a test of general regulatory effectiveness, indicate that the hypothesis of ineffective regulation must be rejected in favor of the model which includes the possibility of a binding rate-of-return constraint for each firm. The results of the second test, a test of individual firm regulatory effectiveness, indicate that rate-of-return regulation has differential effects across firms with the rate-of-return constraint being statistically significant for some firms but not necessarily for all. In addition, regulatory effectiveness is shown to vary significantly across time. These results support the basic contention of this paper that the appropriate model of the regulated firm is one in which λ, the measure of regulatory "tightness" or effectiveness, is allowed to vary across firms. Section 5 contains a comparative discussion of the results obtained in this paper and those of three recent econometric studies of regulatory effectiveness. A brief summary of the paper and several comments concerning implications for future research are offered in Section 6. A technical appendix is included at the end of the paper in which additional theoretical properties of the regulated profit function used in this study are derived.

2. The Averch–Johnson Model of a Regulated Firm

The static equilibrium model of a profit-maximizing firm subject to a regulatory constraint with respect to the maximum earned rate of return on its capital was first analyzed by Averch and Johnson (1962) in a seminal article.[2] In this model the firm is assumed to produce a single output using two inputs, labor and capital. The firm's decisions are influenced by regulation in only one way, namely, in the form of a rate-of-return constraint by which the earned rate of return on capital cannot exceed some allowed "fair" rate of return set by the regulatory commission. This allowed rate of return is assumed to be less than the rate of return the firm might have earned in the absence of regulation – otherwise regulation loses its economic rationale – but greater than the market cost of capital. In the static model it is also assumed that the rate-of-return constraint is always binding on the firm so that the regulated firm always earns exactly the allowed rate of return on its capital. Thus, the static model of the regulated firm rules out regulatory lag as well as any influence on the part of the firm in the determination by the regulatory commission of the allowed rate of return.[3]

Perhaps the most significant implication of the A–J model of the regulated firm is that the firm will choose an inefficient combination of factor inputs in the form of a larger capital–labor ratio than is efficient for the output level that is chosen. This over-capitalization is the direct result of an implicit subsidy to capital in the form of a difference between the allowed rate of return and the market cost of capital, a difference which is assumed to be positive. Another way of explaining this result is to note that capital plays a dual role in the A–J model, namely, as both an input to the production function and as a variable in

[2]In addition to the original A–J article in 1962 and a similar paper written independently by S.H. Wellisz in the same year, the properties of the A–J model have been extended and summarized in two recent studies: Baumol and Klevorick (1970) and Bailey (1973).

[3]The effects of regulatory lag on the A-J model have been studied by Bailey and Coleman (1971), Baumol and Klevorick (1970), Davis (1973), Klevorick (1973), and Sibley and Bailey (1974). The relationship between the regulated firm and the regulatory commission has been explored by Klevorick (1973) and Joskow (1972 and 1973). The effects of uncertainty have been analyzed by Myers (1973) and Holthausen (1974), while the inclusion of alternative objective functions for the regulated firm has been studied by Bailey and Malone (1970).

the rate-of-return constraint. Since profits can be increased by expanding the rate base, it will always pay the firm to use more capital relative to labor than the unregulated (efficient) firm for any given level of output. A somewhat more ambiguous implication of the A–J model is that the output level chosen by the regulated firm will also be inefficient although it will not necessarily be greater than that chosen by the unregulated firm.[4]

We now proceed with a brief review of the theoretical properties of the static A–J model, and in particular the comparative statics properties which will be used in our derivation and discussion of the regulated profit function.

Assume a regulated firm, facing given output and input prices, P, w, and r, respectively, which maximizes profits,

$$P\,F(K,L) - wL - rK, \tag{1}$$

subject to a regulatory constraint on the earned rate of return, s, namely,

$$P\,F(K,L) - wL - sK = 0, \quad \text{for} \quad s > r, \tag{2}$$

and where $Q = F(K,L)$ is a well-behaved neoclassical production function, and L and K are labor and capital inputs, respectively.[5]

The twin assumptions of a given or exogenous output price, P, and an allowed rate of return, s, follows from an examination of the procedures generally used by most state regulatory commissions. In general there are two phases to the process of regulatory review: the determination of a fair rate of return which the regulated firm is allowed to earn on its invested capital or rate base, and the determination of an output price or rate which the firm must use in selling its output.[6] The determination of the allowed rate of return, s, generally includes considerations of the

[4]For discussions of this proposition, see Baumol and Klevorick (1970, pp. 168, 170n, and 176–178) and Bailey (1973, pp. 125–137).

[5]The original A–J model assumed a downward-sloping demand curve, $P = f(Q)$, $f' < 0$, so that both P and Q were assumed endogenous to the firm. As Fuss has pointed out, the duality formulation used in this paper is also compatible with an endogenously determined price as long as a constant elasticity demand curve is assumed since the general result $MR = P(1/(1 - 1/n))$, where MR is marginal revenue and n is the (constant) price elasticity of demand, merely implies a scaling of the output price variable, P. For an example of the use of this assumption, see Spann (1974). However, we shall argue in this study that the essential result of the regulatory process is the determination of an allowed rate of return, s, and an allowed rate or output price, P. Apart from the question of rate structure which is ignored in this study, it would appear, therefore, that the assumption that both s and P are exogenous to the firm is a more tenable one. For further support of this assumption, see Joskow (1973, p. 118).

[6]For more detailed discussions of the process of regulatory review, both formal and informal, see Phillips (1969) and Joskow (1972 and 1974).

firm's current market cost of capital and its past financial and economic performance. Once the regulatory commission decides upon the fair rate of return to be allowed, a second set of calculations are carried out in order to transform the allowed rate of return, s, into an allowed output price or rate, P, which will just permit the firm to earn its allowed rate of return. This second phase of the regulatory hearings usually concentrates upon the determination of the firm's rate base as well as its estimated costs and revenues under the proposed new rate. Since the problem of rate structure, as opposed to level, is assumed away for the purposes of this study, the result is a one-to-one correspondence between the allowed rate of return, s, and the allowed output price, P. Once the output price is set, the regulated firm is required to sell its output to all customers at that price. Thus, the allowed rate of return, s, and the allowed output price, P, are both determined by the regulatory commission and can be assumed to be exogenous to the firm.[7,8]

Equations (1) and (2) can now be simplified somewhat by noting that both sides of both equations can be divided by P and the prices rewritten in terms of normalized prices w', r' and s', equal to w/P, r/P and s/P, respectively. Thus, under the assumption of an exogenous output price, P, the regulated firm can be regarded as maximizing normalized profits,

$$F(K,L) - w'L - r'K, \tag{3}$$

subject to a normalized regulatory constraint

$$F(K,L) - w'L - s'K = 0. \tag{4}$$

[7]Given our assumption concerning the exogenous nature of the output price to the regulated firm and assuming that the regulated firm is a monopolist so that it faces a downward-sloping demand curve, it would appear that the correct specification is one of constrained cost-minimization rather than profit-maximization since output quantity would appear to be determined exogenously. However, the fact that we are dealing at the plant rather than the firm level, that is, with the decision concerning the optimal selection of a new plant; the fact that electric utilities typically buy and sell significant amounts of electricity from other utilities so that own-market demand does not represent total demand for the firm's output; and the fact that the utility has some control over quantity demanded in the form of advertising and substitution effects imply that plant output may not be completely exogenous. Thus, given some doubt as to the correctness of either the cost-minimization or the profit-maximization specification, we choose the latter as a more general specification since it includes cost minimization as a necessary, but not sufficient, condition. For an example of the use of a cost-minimization specification, see Petersen (1975).

[8]While a number of complications, such as the possibility of the firm influencing the determination of s and the possibility of differences between the allowed and the actual rates of return due to uncertainty and regulatory lag, have been ignored in this analysis, it seems reasonable, at least as a first approximation, to assume them away.

This transformation, which is both analytically and econometrically convenient, was developed by Lau (1969c and Chapter I.3) in his discussion of the unit–output–price or UOP profit function, and will be used for the remainder of this study.

The first-order conditions for maximizing (3) subject to the regulatory constraint (4) can be derived in the following form:

$$F_K(1-\lambda) - r' + \lambda s' = 0, \quad (5)$$

$$F_L - w' = 0, \quad (6)$$

$$F - w'L - s'K = 0. \quad (7)$$

Equations (5)–(7) can now be solved simultaneously to obtain the profit-maximizing input demands, $K^*(w',r',s')$ and $L^*(w',r',s')$, and the profit-maximizing values of the Lagrangian multiplier associated with the rate-of-return constraint, $\lambda^*(w',r',s')$.

The second-order conditions for a profit maximum require that

$$-(s'-F_K)^2(1-\lambda)F_{LL} > 0. \quad (8)$$

Since $F_{LL} < 0$, by assumption the second-order conditions imply that $\lambda < 1$.[9]

Several additional properties of λ, the Lagrangian multiplier, can also be noted. The first is that λ is clearly an endogenous variable in the sense that it is selected simultaneously, although implicitly, along with K^* and L^* by the firm. Thus, λ is a function of input prices and, hence, will vary from firm to firm assuming that input prices vary across firms. A second characteristic is that λ must lie in the interval (0,1). The second part of this property, namely, $\lambda < 1$, has already been shown to follow from the second-order conditions. The first part follows from a consideration of the interpretation of λ, namely, λ is the increase in regulated profits resulting from a unit increase in the regulatory constraint (4), i.e., by increasing total revenue through an increase in the normalized rate of return, s'. Assuming that regulation is effective, the result of relaxing the regulatory constraint will always be to permit the firm to earn larger profits so that λ must always be positive. One implication of this result follows immediately by noting that (5) can be solved for λ, yielding

$$\lambda = (r' - F_K)/(s' - F_K). \quad (9)$$

[9]For a more detailed proof, see either Baumol and Klevorick (1970, p. 167) or Bailey (1973, p. 80).

Since $0 < \lambda < 1$ and assuming $s' > r'$, (9) implies that both s' and r' must be greater than F_K so that $s' > r' > F_K$. Finally, the limiting case of $\lambda = 0$ is obviously the case of the ineffectively regulated firm, that is, the case where the rate-of-return constraint (4) is not binding.

Equations (5) and (6) can also be used to obtain the following cost-minimization condition for the regulated firm:

$$F_K/F_L = r''/w', \tag{10}$$

where $r'' = (r' - \lambda s')/(1-\lambda)$. One interpretation of this result is that the regulated firm, faced with (normalized) input prices w' and r' and an allowed rate of return s', acts *as if* it were an unregulated or unconstrained firm facing input prices of w' and r''.[10] Unfortunately, the econometric usefulness of this result is limited since r'' is clearly an endogenous variable due to the influence of λ.

Finally, the comparative statics effects of changes in w', r' and s' on K^*, L^* and λ^* can be derived by means of total differentiation of equations (5)–(7). These results are summarized in Table 1.[11]

Thus, the Averch–Johnson model of the regulated firm, although a rather simple static model, appears to capture the essential elements of the economic behavior of a profit-maximizing firm operating under a regulatory constraint and, in addition, offers a number of testable implications. In particular, tests of regulatory effectiveness can be

TABLE 1
Comparative statics results for the A–J model.

	dw'	dr'	ds'
dK	$-\dfrac{L}{s'-F_K}$	0	$-\dfrac{K}{s'-F_K}$
dL	$\dfrac{LF_{KL}+s'-F_K}{(s'-F_K)F_{LL}}$	0	$\dfrac{KF_{KL}}{(s'-F_K)F_{LL}}$
$d\lambda$	$\dfrac{(1-\lambda)[L(F_{KK}F_{LL}-F_{KL}^2)-(s'-F_K)F_{KL}]}{(s'-F_K)^2F_{LL}}$	$\dfrac{1}{s'-F_K}$	$\dfrac{(1-\lambda)K(F_{KK}F_{LL}-F_{KL}^2)-\lambda(s'-F_K)F_{LL}}{(s'-F_K)^2F_{LL}}$

[10]Equation (10) provides another way of looking at the general A–J result that the regulated firm will always choose to overcapitalize relative to the efficient (competitive) firm since it can be shown that $r'' < r'$ for $s' > r'$. Thus, the regulated firm has an effective or shadow price of capital (services), r'', which is less than the market price of capital (services), r', a condition which provides the incentive for over-capitalization.

[11]Some, although not all, of these results have also been derived by Baumol and Klevorick (1970) and Bailey (1973, Ch. 8).

carried out on λ since $\lambda = 0$ implies ineffective regulation while $\lambda > 0$ implies a binding regulatory constraint. However, the design of an econometric model for carrying out such tests must deal explicitly with two essential characteristics of λ. The first is that λ is an endogenous variable and in general will vary across firms and, hence, across observations. The second characteristic is that λ is intrinsically unobservable so that an estimation procedure must be used which does not require observations on λ. As we shall demonstrate in the next section, the profit function approach offers an econometric framework which enables both of these problems to be handled satisfactorily.

3. The Profit Function for a Regulated Firm

Since past studies of regulatory effectiveness have suffered from a number of potentially serious problems, primarily with respect to the proper specification of the model, there is need for further econometric analysis using the Averch–Johnson model. The profit function offers an ideal starting point for such an analysis since, as we have shown, the regulated firm can be regarded as facing given prices for both inputs and output and can be assumed to be a profit-maximizing firm. However, the rate-of-return constraint must be explicitly imbedded within the regulated profit function if misspecification problems are to be avoided. The purpose of this section, therefore, is to derive a general regulated profit function which takes explicit account of the rate-of-return constraint.

In general, the unregulated profit function expresses maximum profits as a function of output and input prices for the case of the competitive profit-maximizing firm as can be seen by substituting the profit-maximizing input demands, $K^*(P,w,r)$ and $L^*(P,w,r)$, into the definition of (unregulated) profits [see equation (1)] yielding

$$\pi^*(P,w,r) = PF[K^*(P,w,r),L^*(P,w,r)] - wL^*(P,w,r) - rK^*(P,w,r). \tag{11}$$

A somewhat simplier version of (11), the so-called UOP or normalized profit function developed by Lau in which the arguments are normalized input prices, w/P and r/P, respectively, can be similarly derived,

$$\pi^*(w',r') = F[K^*(w',r'),L^*(w',r')] - w'L^*(w',r') - r'K^*(w',r'), \tag{12}$$

and will be used in the following derivation.[12]

One of the most important properties of profit functions is contained in Hotelling's lemma which states that the partial derivatives of the profit function with respect to each of the input prices are equal to the negative of the respective profit-maximizing input demands.[13] In terms of the UOP profit function, $\pi^*(w',r')$, we have

$$\partial\pi/\partial w' = -L^*(w',r') \quad \text{and} \quad \partial\pi/\partial r' = -K^*(w',r'). \tag{13}$$

From an econometric point of view, this result is extremely useful since it permits one to write a profit function directly and derive estimable input demand equations from it, either in terms of input quantities or relative shares, without explicitly specifying the underlying production technology. Duality theory assures that the correspondence between the specified profit function and the underlying production function is unique, and since the problem of deriving the associated profit function from a given production function is often intractable, profit functions offer a powerful tool for the estimation of production technologies.

In the case of a regulated firm, that is, a profit-maximizing firm subject to a rate-of-return constraint, Hotelling's lemma for the unregulated or unconstrained firm, (13), does not hold. However, a revised version of this result does hold so that the profit function approach can be extended to the case of a regulated profit-maximizing firm by means of what we shall call the regulated profit function. We now proceed to derive this revised version of Hotelling's lemma for the special case of a profit-maximizing firm subject to a rate-of-return constraint.[14]

The Lagrangian expression for the constrained optimization problem of the regulated firm can be written as

$$PQ - wL - rK - \lambda(PQ - wL - sK). \tag{14}$$

Dividing (14) by the output price P yields an equivalent expression in terms of normalized input prices w', r' and s', which we will call $\mathcal{L}$, so

[12]It should be noted that under the assumption of a given output price, the UOP profit function has the same duality properties associated with it as the general profit function (11). In particular, Hotelling's lemma concerning the partial derivatives of the profit function with respect to the input prices remains valid as does the one-to-one correspondence between profit function and production function parameters, e.g., returns to scale and substitution characteristics.

[13]The original statement of this proposition seems to be Hotelling (1932). More recent proofs have been offered by Shepherd (1953), McFadden (1966) and Lau (1969c).

[14]This proof was suggested by M. Fuss. For an alternative proof of this proposition, see Section 7.

that

$$\mathscr{L} = Q - w'L - r'K - \lambda(Q - w'L - s'K). \tag{15}$$

The assumption of an exogenous or given output price P guarantees that the optimization solution of (15) in terms of w', r' and s' will be equivalent to that of (14) in terms of P, w, r and s. Letting (*) denote optimal values, we further note that

$$\mathscr{L}^* = Q^* - w'L^* - r'K^* = \pi^*, \tag{16}$$

since the expression $\lambda(Q - w'L - s'K)$ is always equal to zero when evaluated at optimal values for λ, K and L. Thus, the partial derivatives of the UOP profit function π^* can be evaluated by differentiating (15) with respect to the normalized input prices, w', r' and s', and assuming that all choice variables take on their optimal values, that is, that the first-order conditions (5)–(7) hold.

For example, differentiating (15) with respect to w' we obtain

$$\frac{\partial \mathscr{L}}{\partial w'} = -(1-\lambda)L + \frac{\partial K}{\partial w'}[F_K(1-\lambda) - r' + \lambda s'] + \frac{\partial L}{\partial w'}[(1-\lambda)(F_L - w')] - \frac{\partial \lambda}{\partial w'}[Q - w'L - s'K].$$

The first-order conditions for profit maximization, (5)–(7) imply that the expressions inside the square brackets vanish so that, noting (16), we obtain

$$\partial\pi^*/\partial w' = \partial\mathscr{L}^*/\partial w' = -(1-\lambda^*)L^*. \tag{17}$$

Similarly, differentiating with respect to r' and s', respectively, yields

$$\partial\pi^*/\partial r' = \partial\mathscr{L}^*/\partial r' = -K^*, \tag{18}$$

and

$$\partial\pi^*/\partial s' = \partial\mathscr{L}^*/\partial s' = \lambda^* K^*. \tag{19}$$

Equations (17)–(19) constitute the revised version of Hotelling's lemma for the special case of a regulated profit-maximizing firm.

A general reduced-form econometric scheme for the regulated firm can now be derived from (17)–(19) by eliminating λ, since λ is an unobservable variable. At the same time, this scheme permits estimates of λ to be retrieved since λ will be a function of the estimated parameters as well as of the observed prices. Denoting $\partial\pi^*/\partial w'$, $\partial\pi^*/\partial r'$ and $\partial\pi^*/\partial s'$ as $\pi_{w'}$, $\pi_{r'}$ and $\pi_{s'}$, respectively, we note that $(1-\lambda)$ can be written as $(\pi_{r'} + \pi_{s'})/\pi_{r'}$.[15] Thus, including the regulated UOP profit function as a separate equation, we can express (17)–(19) equivalently as[16]

$$\begin{aligned} \pi &= \pi(w',r',s'), \\ -L &= \pi_{w'} \cdot \pi_{r'}/(\pi_{r'} + \pi_{s'}), \\ -K &= \pi_{r'}. \end{aligned} \tag{20}$$

The system of equations in (20) is a set of reduced-form equations in the observable prices w', r' and s', and contains all of the restrictions implied by the model of the regulated firm, (3) and (4). Under the assumption that the output price P can be assumed to be exogenous to the regulated firm, (20) can be used to test for regulatory effectiveness within a framework which implicitly treats λ as an endogenous variable. This last property means that (20) offers an improved econometric specification over that used by previous studies since it does not restrict λ to be a constant. In addition, estimates of λ for each observation can be computed from the estimated parameters of the regulated profit function by noting from (18) and (19) that

$$\lambda = -\pi_{s'}/\pi_{r'}, \tag{21}$$

so that variations in regulatory effectiveness across firms can also be tested for, a test which is impossible within the framework generally used by previous studies.

4. Statistical Tests of Regulatory Effectiveness

With (20) as the general econometric model, we now proceed to a specific functional specification of the *ex ante* technology of a regulated electric utility in order to derive specific statistical tests of regulatory effectiveness.

Since conventional electricity generation requires three inputs – fuel, capital and labor – we first extend (20) to the case of three inputs where F represents fuel inputs and f' represents the normalized price of fuel.

[15]We now drop the starred notation for expositional convenience. Unless otherwise noted, all quantities will henceforth refer to optimal quantities.

[16]The curvature properties of the regulated profit function, $\pi = \pi(w',r',s')$, are derived in Section 7.2.

This extension is straightforward, and results in the following model:

$$\begin{aligned} \pi &= \pi(f',w',r',s'), \\ -F &= \pi_{f'}\cdot\pi_{r'}/(\pi_{r'}+\pi_{s'}), \\ -L &= \pi_{w'}\cdot\pi_{r'}/(\pi_{r'}+\pi_{s'}), \\ -K &= \pi_{r'}. \end{aligned} \tag{22}$$

For the purposes of this study, we adopt a quadratic specification of the regulated profit function, a specification which can be interpreted as a second-order Taylor series approximation to the true underlying profit function.[17] Thus, we assume the following regulated profit function where all input prices have been normalized in terms of the output price,[18]

$$\begin{aligned} \pi^{\text{reg}} = {} & \beta_0+\beta_1 f+\beta_2 w+\beta_3 r+\beta_4 s+(\beta_5/2)f^2+(\beta_6/2)w^2 \\ & +(\beta_7/2)r^2+(\beta_8/2)s^2+\beta_9 f\cdot w+\beta_{10} f\cdot r+\beta_{11} f\cdot s \\ & +\beta_{12} w\cdot r+\beta_{13} w\cdot s+\beta_{14} r\cdot s. \end{aligned} \tag{23}$$

Rather than writing out the full system, equivalent to (22), we simply note that the partial derivatives of (23) are

$$\begin{aligned} \pi_f &= \beta_1+\beta_5 f+\beta_9 w+\beta_{10} r+\beta_{11} s, \\ \pi_w &= \beta_2+\beta_9 f+\beta_6 w+\beta_{12} r+\beta_{13} s, \\ \pi_r &= \beta_3+\beta_{10} f+\beta_{12} w+\beta_7 r+\beta_{14} s, \\ \pi_s &= \beta_4+\beta_{11} f+\beta_{13} w+\beta_{14} r+\beta_8 s, \end{aligned} \tag{24}$$

and that the substitution of (23) and (24) into (22), and the addition of classical additive disturbance terms in each of the four equations, yields an estimable, albeit extremely non-linear, system of reduced-form equations.

Consistent estimates of the parameters of this system were computed using the maximum likelihood estimation scheme suggested, for exam-

[17]For the purposes of this study, the quadratic specification has two advantages over an alternative translog specification. The first is that it avoids the local versus global convexity problem associated with the translog specification. Although this is not a major problem it does require that an estimated translog function be tested for convexity for each set of observations since global convexity is not assured. See Lau (Appendix A.4) for further discussion of this problem. A second problem with the relative share demand equations derived from a translog specification is that they are not independent, since the relative shares must sum to one, and so one equation must be dropped from the estimation scheme with a resulting loss in efficiency. This problem is avoided with a quadratic specification.

[18]For the remainder of this paper, the primed notation will be dropped. Unless otherwise noted, all variables refer to normalized (in terms of output price) variables.

ple, by Malinvaud (1970). This scheme uses Zellner's minimum distance estimator and the Gauss–Newton computational method. The specific program used was an iterative non-linear program whose convergence criteria depended upon both the largest change in parameter values associated with the last iteration, and the difference between the transformed residual covariance matrix and an identity matrix of order equal to the number of equations being estimated.[19] For most cases in this paper, convergence criteria of 0.01 or 1.0% were used.

The null hypothesis that regulation is ineffective can be tested as a nested hypothesis by noting that the case of a non-binding constraint on the regulated firm would imply a special case of this system, namely,

$$\beta_4 = \beta_8 = \beta_{11} = \beta_{13} = \beta_{14} = 0. \tag{25}$$

This result can easily be seen by noting that the equivalent system of equations for an ineffectively regulated profit-maximizing firm would be

$$\begin{aligned} \pi &= \beta_0 + \beta_1 f + \beta_2 w + \beta_3 r + (\beta_5/2)f^2 + (\beta_6/2)w^2 \\ &\quad + (\beta_7/2)r^2 + \beta_9 f\cdot w + \beta_{10} f\cdot r + \beta_{12} w\cdot r, \\ \pi_f &= \beta_1 + \beta_5 f + \beta_9 w + \beta_{10} r, \\ \pi_w &= \beta_2 + \beta_9 f + \beta_6 w + \beta_{12} r, \\ \pi_r &= \beta_3 + \beta_{10} f + \beta_{12} w + \beta_7 r, \end{aligned} \tag{26}$$

which is equivalent to the system implied by (23) and (24) under the additional restriction of (25). Thus, the overall effectiveness of rate-of-return regulation can be tested using the likelihood ratio test resulting from the estimation of (22)–(24) with and without (25) imposed.

This test of regulatory effectiveness is actually a test that λ is simultaneously equal to zero for all firms in the sample, that is, that regulation is ineffective for all firms simultaneously rather than only for some. This can be seen by noting that (25) implies that $\pi_s = 0$ which, from (21), implies that $\lambda = 0$ for all firms. Thus, this test appears overly restrictive since regulatory ineffectiveness is more likely to mean that regulation is ineffective for some firms but not necessarily for all. This is especially likely to be the case if the regulated firms come from different states with significant variations in regulatory procedures among states. Thus, it is desirable that a test for regulatory effectiveness be used which permits the regulatory constraint to be binding on some firms but

[19]For other examples of the use of this estimator, see Atkinson and Halvorsen (1976) and Christensen, Jorgenson and Lau (1975).

not necessarily on all. Accordingly, a second test of regulatory effectiveness will be derived, one which applies to individual firms.

The individual estimates of the λ's can be tested against the null hypothesis $\lambda_i = 0$, $i = 1,...,n$, by estimating the system of equations (22) without any additional restrictions. Equation (21) can then be used to obtain estimates of the λ_i's, and under the assumption that λ is asymptotically normally distributed, confidence intervals can be estimated and used to test the null hypothesis $\lambda_i = 0$. In addition to being consistent with the theoretical model, in the sense that λ is allowed to vary across firms, this test also permits one to look for factors such as firm size, state location and method of evaluation (original cost versus fair value) which may influence the effectiveness of rate-of-return regulation across firms.[20]

4.1. *Data*

The data used in this study was drawn primarily from a previously collected sample of 150 new privately-owned steam-electric generating plants constructed between 1947 and 1965.[21] Of the 150 plants in the original sample, 33 were deleted on the grounds that they were located in states with no statewide regulation, 2 on the grounds that they were jointly owned by two or more firms, and 1 on the grounds that it appeared to be primarily an industrial plant, leaving 114 plants for the purpose of this study. In many cases, missing data meant that somewhat less than 114 observations were available.

Since the Averch–Johnson model is a long-run equilibrium model with an emphasis upon the optimal quantity of capital for the regulated firm, the proper specification for testing for regulatory effectiveness is an *ex ante* profit function in which capital is assumed to be a variable input. As a result all prices must be treated as anticipated or expected prices on the part of the firm installing the new plant since the actual design installed was presumably selected from a number of feasible technologies on the basis of future expected input and output prices. The

[20]In addition, the duality between profit function and production functions implies that several measures of the underlying technology, such as elasticities of substitution and the degree and bias of the scale effect, could also be estimated from these results as functions of the estimated coefficients and the observed input prices. Since the focus of this study is on regulatory effectiveness, however, these calculations were not carried out. For an example of this alternative focus, see Atkinson and Halvorsen (1976).

[21]For a more detailed description of this data, see Cowing (1974).

data used in this study consisted of design data with respect to both output and input quantities, and input price data for several years prior to the installation of the plants. Thus, it was assumed that this prior years input price data was a reasonable proxy for the expected or anticipated input prices which were presumably used in the *ex ante* selection of the (optimal) plant design.

The cost of capital, r, was measured by the bond rate for each firm's bond issue immediately preceding construction of the new plant. Since detailed physical depreciation information was not available, the rate of depreciation was assumed to be constant across firms and, hence, was not included within the cost of capital measure. The allowed rate of return, s, was measured by the actual earned rate of return on total capital for each firm for the year preceding plant installation. Under the assumption of static equilibrium, the regulated firm will always earn precisely the allowed rate of return so that the earned rate of return can be used as a proxy for the allowed rate of return.[22] In the cases of both the price of fuel, f, and the wage rate, w, two-year regional averages for the two years prior to plant installation were used. The wage rate was based on the regional wage rates of production workers in the electric utility industry, while the price of fuel was based on a regional average for the dominant fuel used by the new plant. Fuel input, F, was measured by multiplying the expected output of the plant by the design heat rate, a measure of plant thermal efficiency expressed in terms of BTU's of fuel input per net kilowatt–hour of electrical output. The design labor force in terms of total employees was used as a measure of labor input, L. Finally, capital input, K, was measured by the total expected cost of constructing the plant.

The expected output of the plant, Q, was measured as the product of the designed capacity in kilowatts, the designed load factor – a measure of expected average capacity utilization – and the number of hours per year. Since data on the price of output are not available at the plant level, the price of output, P, was measured as the average revenue from electricity sales for the firm owning the plant for the year preceding plant installation. Finally, profits were computed as the difference between total revenues and total cost. Both input prices and profits were normalized in terms of the price of output.

The effects of technical change were taken into account by estimating the system of equations (22) separately for each of the four tech-

[22]This approach was also used by Spann (1974). A modified version with respect to the return on equity capital was used by Petersen (1975).

nological epochs used by Dhrymes and Kurz (1964): 1947–50, 1951–54, 1955–59 and 1960–65. This procedure implicitly assumed that there was no intra-epoch technical change. Given the small number of years within each of these epochs, this is probably not an unreasonable procedure even in light of the fact that technical change in the electric utility industry during the 1950's and early 1960's appears to have been substantial.

4.2. *Statistical Results*[23]

The results of the first test, the test of general regulatory effectiveness, are summarized in Table 2 for three periods: 1947–50, 1955–59 and 1960–65.[24] Model II is the unrestricted model (22) which, as we have shown, is the general model of a regulated firm. Model I is (22) with the restrictions in (25), that is, $\beta_4 = \beta_8 = \beta_{11} = \beta_{13} = \beta_{14} = 0$, imposed. Model III contains the restriction $\beta_7 = 0$, a restriction which can be derived

TABLE 2
Tests of general regulatory effectiveness.[a]

	Number of observations	Log of likelihood function	Δl	$-2(\Delta l)$
(a) 1947–50				
Model I	21	−132.6	−21.2	42.4
Model II	21	−111.4	–	–
Model III	21	−135.9	−24.5	49.0
(b) 1955–59				
Model I	26	−201.0	−17.1	34.2
Model II	26	−183.9	–	–
Model III	26	−185.8	−1.9	3.8
(c) 1960–65				
Model I	23	−176.8	−14.0	28.0
Model II	23	−162.8	–	–
Model III	23	−163.9	−1.1	2.2

[a]Model I imposes the restriction $\beta_4 = \beta_8 = \beta_{11} = \beta_{13} = \beta_{14} = 0$. Model II is the unrestricted version, while Model III imposes the restriction $\beta_7 = 0$.

[23]In order to reduce the non-linearity, a three-equation version of (22) using F/L rather than $-F$ and $-L$ separately, was estimated.

[24]Convergence problems with the iterative non-linear estimation program used for the period 1951–54 forced us to discard the results for this time period so that only three time periods are reported on in the remainder of the study.

from the A–J model as shown below. As is well-known, the product of −2 and the log of the ratio of the unrestricted and restricted values of the likelihood function has an asymptotic chi-squared distribution with q degrees of freedom where q is the number of imposed restrictions.[25] Thus, general regulatory effectiveness can be tested by estimating (22) with and without (25) imposed, and then using a likelihood ratio test. The log of the likelihood ratio is shown under the column labeled Δl in Table 2, while the last column gives the values of the corresponding chi-squared statistic, that is, $-2\Delta l$.

A comparison of the results for Models I and II for each of the three time periods shown indicates that the model of the unregulated or ineffectively regulated firm, that is, the null hypothesis that $\lambda = 0$, can be rejected in each case, since the critical chi-squared value for 5 degrees of freedom at the 0.01 significance level is 15.09. Thus, we conclude that rate-of-return regulation in the case of electric utilities for the three time periods shown was generally effective. Although it is quite possible that regulation is not simultaneously binding on all firms, our results indicate that regulation is effective on enough firms to produce an overall picture of general regulatory effectiveness.

An additional test of general regulatory effectiveness is possible within the framework of (22). This additional test follows from noting that the A–J model of the regulated firm implies that $\pi_{rr} = 0$, a result which is derived in the appendix but which is also obvious from the results in Table 1 by noting that $\pi_{rr} = \mathrm{d}K/\mathrm{d}r = 0$. Inspection of the third equation in (24) shows that this implies

$$\beta_7 = 0, \tag{27}$$

a result which represents a testable implication of the theory of the regulated firm. The results of this test are shown in Table 2 under Model III, and indicate that the restriction (27) cannot be rejected in favor of the unrestricted model (22) for two of the three time periods. Only for the period 1947–50 does Model III do significantly worse than the unrestricted case, Model II. Thus, there is at least some evidence that the data are not inconsistent with this last property of the A–J model, a result which further supports the conclusion of general regulatory effectiveness within an A–J framework.

The estimated coefficients for Model II, as well as the standard errors

[25]See Wilks, (1963). For a proof that the iterative estimation technique used here yields an estimated likelihood ratio having an asymptotic chi-squared distribution, see Zellner (1962).

of regression for each of the three equations, are shown in Table 3 for each of the three time periods. While the results for the first two time periods suggest a reasonable performance in terms of the number of significant coefficients, the results for the 1960–65 period indicate that only three of the coefficients, β_7, β_8 and β_{14}, are significant at the 0.1 level using a two-tailed test. Another indication of performance is available by noting that the A–J model of the regulated firm implies not only $\beta_7 = 0$ but also the following restrictions:

$$\begin{aligned} &\text{(i)} \quad \pi_{ss} < 0 \quad \text{which implies that} \quad \beta_8 < 0, \\ &\text{(ii)} \quad \pi_{wr} > 0 \quad \text{which implies that} \quad \beta_{12} > 0, \\ &\text{(iii)} \quad \pi_{rs} > 0 \quad \text{which implies that} \quad \beta_{14} > 0. \end{aligned} \tag{28}$$

Unfortunately, Table 3 shows that these restrictions cannot generally be supported by our statistical results. For example, the hypothesis $\beta_7 = 0$

TABLE 3
Estimated coefficients for Model II.[a,b]

Regression coefficient	1947–50	1955–59	1960–65
β_0	2.237 (3.37)	−1.346 (−1.98)	2.286 (0.89)
β_1	−0.052 (−0.58)	−4.434 (−6.01)	1.701 (0.61)
β_2	0.104 (0.73)	−3.809 (−5.97)	0.730 (0.60)
β_3	−1.895 (−3.29)	0.301 (0.63)	−2.961 (−1.50)
β_4	0.033 (0.11)	3.394 (5.94)	0.270 (0.30)
β_5	1.116 (2.32)	1.679 (3.98)	−0.922 (−0.61)
β_6	0.742 (2.43)	1.914 (3.59)	−0.179 (−0.58)
β_7	2.445 (3.26)	1.402 (1.78)	2.884 (1.99)
β_8	0.010 (0.08)	−1.440 (−4.62)	1.187 (2.04)
β_9	0.857 (2.49)	1.669 (3.76)	−0.391 (−0.60)
β_{10}	−1.388 (−2.49)	−1.168 (−2.09)	0.013 (0.17)
β_{11}	0.176 (2.34)	1.056 (3.31)	0.040 (0.59)
β_{12}	−1.148 (−2.53)	−1.135 (−2.07)	0.010 (0.24)
β_{13}	0.117 (2.10)	0.758 (2.24)	0.017 (0.63)
β_{14}	−0.188 (−0.78)	−0.305 (−0.95)	−1.576 (−1.84)
	Standard Error of Regression		
Equation			
π	0.441	0.597	0.842
$-K$	1.245	2.329	3.029
L/F	0.166	0.452	0.114

[a]Model II is the unrestricted (regulated) model.
[b]Numbers shown in parentheses are t-values.

is rejected in 2 of the 3 cases shown,[26] as is the hypothesis $\beta_8 < 0$, while the hypotheses $\beta_{12} > 0$ and $\beta_{14} > 0$ are rejected in all 3 cases. Thus, the model does not perform well in terms of these implied restrictions of the general A–J model.

The combined results of Tables 2 and 3 appear to indicate that some form of regulatory framework is effective but perhaps not one with the

TABLE 4
Tests for regulatory effectiveness across firms for Model II.[a,b]

Number of observation	Estimates of λ_i 1947–50	1955–59	1960–65
1	0.388 (0.73)	0.430 (4.31)	0.380 (3.23)
2	0.078 (0.78)	0.208 (2.59)	0.492 (5.45)
3	0.090 (1.08)	0.177 (2.37)	0.673 (4.96)
4	−0.003 (−0.01)	0.414 (1.60)	1.063 (2.07)
5	0.363 (0.36)	−0.149 (−0.52)	0.497 (5.44)
6	0.040 (0.34)	−0.149 (−0.52)	0.618 (6.54)
7	0.074 (1.13)	0.713 (3.72)	0.679 (4.88)
8	0.037 (0.31)	0.698 (1.79)	1.351 (0.60)
9	0.049 (0.47)	0.859 (2.91)	0.153 (0.17)
10	0.047 (0.42)	−0.805 (−2.09)	0.313 (2.08)
11	0.031 (0.20)	−46.80 (−0.01)	0.316 (2.30)
12	0.050 (0.51)	0.761 (3.40)	0.465 (4.85)
13	0.041 (0.29)	0.575 (4.07)	0.558 (5.75)
14	0.096 (2.44)	0.531 (0.83)	0.213 (0.87)
15	0.091 (1.28)	−12.31 (−0.16)	−0.049 (−0.07)
16	0.064 (0.78)	1.346 (3.37)	0.494 (5.36)
17	0.064 (0.68)	0.222 (1.49)	0.111 (0.41)
18	0.062 (0.79)	0.660 (4.01)	0.713 (4.86)
19	0.095 (0.55)	2.634 (3.07)	0.870 (3.38)
20	0.096 (1.58)	−0.741 (−0.77)	0.905 (3.30)
21	0.062 (0.78)	0.900 (3.55)	0.533 (6.14)
22		−1.108 (−2.53)	0.392 (3.60)
23		−0.103 (−0.76)	−6.373 (−0.04)
24		−0.013 (−0.07)	
25		−1.122 (−3.44)	
26		−2.607 (−2.65)	

[a]Model II is the unrestricted (regulated) firm.
[b]Numbers shown in parentheses are t-values.

[26]This statement may appear to contradict the earlier decision that $\beta_7 = 0$ cannot be rejected for 2 of the 3 periods (Table 2). However such is not necessarily the case since the two decisions are based on different test statistics. The decision in Table 2 is based on the likelihood ratio test statistic while the one in Table 3 is based on the Wald test statistic. Although the two statistics have identical asymptotic distributions; in multivariate regression models their finite numerical values may differ and thus conflicting decisions can arise. For a detailed discussion of conflicts of this type, see Berndt and Savin (1977).

exact structure implied by the static Averch–Johnson model used in this study. However, since this model ignores such features as regulatory lag and the possibility of interdependencies between the regulatory commission and the regulated firm, this result is not too surprising. No doubt a more refined version of this model which incorporated some of these additional features of the regulatory process would perform better. However, even within the simple regulatory framework assumed here, the main conclusion would appear to be one of general regulatory effectiveness.

A second test of regulatory effectiveness for individual firms is possible by testing the null hypothesis $\lambda_i = 0$ against the alternative hypothesis $\lambda_i > 0$. Assuming an asymptotically normal distribution for λ and using approximate confidence intervals, Table 4 shows the estimated λ's and the associated t-values. Note that there is one λ for each observation in each sample. Using a one-tailed test, the results of this test can be summarized as follows in Table 5.

TABLE 5
Results of one-tailed test on $\hat{\lambda}_i$.

Period	Results[a]
1947–50	1 of 21 firms significant
1955–59	12 of 26 firms significant
1960–65	17 of 23 firms significant

[a]At the 0.05 significance level.

Thus, the results appear to indicate considerable variation in regulatory effectiveness both across firms and across time. For example, our results indicate that for the period 1947–50 all but one of the firms in our sample were not constrained by rate-of-return regulation so that such regulation appears to have been generally ineffective for that period. On the other hand, regulation of electric utilities appears to have been generally effective during the period 1960–65 since most of our observations show significant λ's.

It is interesting to compare our results, based upon a slightly revised version of the general Averch–Johnson model of the regulated firm, with both the model and results in a recent study by Joskow (1974). In that study, Joskow argues for a model of differentially effective rate-of-return regulation across time, similar to our contention concerning variations in λ, from periods of ineffective regulation to periods when the rate-of-return constraint is binding. However, in contrast to the A–J model which assumes a primary regulatory focus upon the earned rate of return, Joskow argues that regulatory commissions appear to be more interested in keeping nominal output prices from rising and that they will allow virtually any earned rate of return as long as the regulated firm does not request a rate increase. This alternative and quite plausible view of the regulatory process implies two phases to regulation: a passive phase during which nominal prices remain constant, or perhaps are even voluntarily reduced somewhat by the firm, and any earned rate of return is allowed; and an active phase in which cost pressures and a resulting low or falling earned rate of return force the firm to file a request for a rate increase. Thus, according to Joskow:[27]

> There is no "allowed" rate of return that regulatory commissions are continually monitoring and at some specified point enforcing......Regulatory reviews are......initiated by requests for nominal price increases and not by the drift of rates of return above some imaginary "allowed" level.

Joskow's model, therefore, is clearly more institutionally oriented and presents an alternative specification of the regulatory process to that of the Averch–Johnson framework.

While Joskow presents several pieces of evidence which appear to be consistent with the implications of his alternative view of the regulatory process, we shall discuss only one of these in this study, namely, variations over time in the annual number of formal rate-of-return reviews presented to state regulatory commissions. Joskow's model implies that formal regulatory proceedings will be initiated primarily by the firm as the result of cost pressures, due to rising input prices which are not offset either by scale economies or technical change, which cause an unsatisfactory earned rate of return and thus force the firm to file for a rate increase in order to raise its earned rate of return in the future. Thus, one would expect a positive correlation between periods of rising costs and the number of formal rate hearings, an implication which is confirmed by the data on the number of formal hearings per year over the period 1949–1972, shown in his Table I. We shall argue

[27]Joskow (1974, pp. 298–299).

that although this data is consistent with Joskow's model, it is also consistent with the findings in our study and, hence, cannot be used as evidence for rejecting the A–J model. Indeed, it would appear that this data, at least in the form presented in Joskow's study, may not be capable of distinguishing between the two alternative models of the regulatory process.

Joskow's data reveals three periods of rather substantial regulatory activity, 1949–53, 1958–60, and 1969–72, all of which he argues were associated with rising input prices and increasing average costs. In contrast, the period 1961–68 shows much lower levels of formal regulatory activity, presumably because of more stable cost conditions. Thus, there appears to be a strong relationship between the strength of upward cost pressures on regulated utilities and the amount of formal regulatory activity, a finding which, as we have noted, is consistent with the implications of Joskow's model. However, these same results are also consistent with the results of our study, a study which uses a modified A–J model of the regulatory process.

Our results, as summarized in Table 5, indicate that the rate-of-return constraint was not generally binding over the period 1947–50 but was generally effective over the period 1960–65.[28] Since the rate-of-return constraint in our model will be non-binding in periods when the earned rate of return is less than the allowed rate of return, due either to an inadequate rate level or to rising costs, it is clear that we would expect an increased number of formal rate increases for the period 1974–50 on the basis of our results.[29] Conversely, our results for the period 1960–65 indicate that the rate-of-return constraint was generally binding which implies that most utilities were able to earn the allowed rate of return. Thus, we would expect a relatively small amount of formal regulatory activity during this period, a result which is confirmed by Joskow's data. As a result, it would appear that, at a minimum, Joskow's data offers further evidence supporting the model used in our study, and, in addition, implies that his conclusion of the general inappropriateness of the A–J model is not substantiated since his data do not appear to offer grounds for rejecting the A–J model, at least in the form we have used. Thus, it would appear that further testing of the two alternative models of rate-of-return regulation is warrented.

[28]The intermediate period, 1955–1959, shows more mixed results with about one-half of the observations exhibiting a binding regulatory constraint.

[29]For further discussion of the A–J regulatory constraint and the conditions under which it will be binding, see Bailey (1973).

5. Comparison with Previous Studies

The results of the above tests indicate generally that the regulatory constraint is binding on the profit-maximizing regulated firm, although with substantial variation across both firms and time. Thus, the Averch–Johnson model of the regulated firm may have considerable empirical validity, a conclusion which has recently been suggested by several other studies. Although there have been a number of econometric studies of the electric utility industry in the past two decades, none of these early studies took explicit account of the regulated nature of the industry.[30] It is only recently that three studies have appeared which have attempted, either directly or indirectly, to test for the effectiveness of rate-of-return regulation: Spann (1974), Courville (1974) and Petersen (1975).

Spann's study (1974) was the first published attempt at direct estimation of λ, the Lagrangian multiplier in the A–J model, which, as we have shown, is a measure of regulatory effectiveness. Spann assumed a translog production function with three inputs, fuel, labor and capital, and used the first-order equilibrium conditions for a profit-maximizing monopoly subject to a rate-of-return constraint to derive the A–J restrictions on the factor share equations. His model consisted of the following two factor share equations:

$$u_K = b_1 + b_2 \ln K + b_3 \ln F + b_4 \ln L + \lambda Z, \tag{29}$$

and

$$u_F = b_5 + b_6 \ln K + b_7 \ln F + b_8 \ln L, \tag{30}$$

where F, L and K are fuel, labor and capital inputs, respectively; u_K is capital's share of gross revenue, rK/PQ; u_F is fuel's share of gross revenue, qF/PQ; $Z = sK/PQ$; and r, q and s are the cost of capital, the price of fuel and the allowed rate of return, respectively. In addition, Spann showed that the following inter-equation restriction was also implied by the model, namely,

$$(1-\lambda)b_6 = b_3. \tag{31}$$

The specific estimation scheme used by Spann consisted of a non-linear search procedure on λ using (29) and (30) in which the objective function was to minimize the total sum of squared errors around the first two equations. This set of simultaneous equations was estimated under the alternative restrictions of $(1-\lambda)b_6 = b_3$ and $b_6 = b_3$, i.e., $\lambda \neq 0$ and

[30] See Cowing (1970) for a discussion of these early studies of the electric utility industry.

$\lambda = 0$, respectively. A chi-squared statistic based on the two sums of squared errors, i.e., under the null and alternative hypotheses, respectively, was used to test the null hypothesis $\lambda = 0$. Two sets of data were used: plant data on 35 new steam electric plants constructed between 1959 and 1963, and firm data on 24 large electric utilities for 1963. These periods were used since it was felt that the relatively stable cost and demand conditions which generally existed during these periods were more likely to result in observations representing long-run equilibrium points.

Spann was able to reject the null hypothesis of an insignificant A–J effect at the 0.01 significance level using two separate measures of capital, megawatt capacity and book value of assets, and for both sets of data, plant and firm. His estimates of λ were in the range 0.5 to 0.7 and, thus, were consistent with the theoretical model.[31] Spann concluded that the Averch–Johnson restrictions with respect to the rate-of-return constraint were significant, that is, that regulation was effective, at least for the time period included within his study.

Courville (1974) used a rather different approach to test for regulatory effectiveness within an Averch–Johnson framework. He first estimated the production function for steam-electric generation and then used the estimated parameters to compute the marginal rate of technical substitution between labor and capital inputs. Since one implication of the A–J model is that the ratio of the marginal products of capital and labor is less than the ratio of their respective input (market) prices, Courville derived a statistical test of regulatory effectiveness based on the difference between these two ratios.

For his production function, Courville assumed a Cobb–Douglas specification, namely,

$$\ln Q = \ln A + a \ln K + b \ln F + c \ln L + dU + eC, \qquad (32)$$

where Q is output; K, L and F are capital, labor and fuel inputs, respectively; U is a measure of capacity utilization; and C is capacity. Since the ratio of the marginal products of capital and fuel using the Cobb–Douglas specification is equivalent to (aF/bK), regulatory effectiveness can be tested by testing the null hypothesis,

$$aF/K - bP_K/P_F = 0, \qquad (33)$$

[31]Spann also attempted a test of the assumption of profit maximization using this model but met with mixed results. See his Table 3 [Spann (1974, p. 49)].

against the alternative hypothesis,

$$aF/K - bP_K/P_F < 0, \tag{34}$$

where P_K and P_F are the prices of capital and fuel inputs, respectively.

Courville's data consisted of first-year-of-operation observations on 110 new steam-electric generating plants for the four periods 1948–50, 1951–55, 1956–59 and 1960–66. Since the coefficient for labor was not statistically significant, the results of the study were based on a revised version of (32) with ln L dropped. This equation was estimated separately for each of the three time periods using two different measures of capital, deflated and undeflated.

Courville was able to reject the null hypothesis (33) in favor of (34) for 105 of the 110 observations at the 0.05 significance level using the undeflated measure of capital, and for 107 of the 110 observations at the 0.05 significance level using the deflated measure of capital. Revised tests in which he attempted to account for bias due to biased estimates of both expected output and fuel expenses resulted in somewhat smaller numbers of rejection of the null hypothesis, but the overall conclusion was sustained, namely, the existence of significant overcapitalization in the electric utility industry presumably due to regulatory effectiveness. To gain some idea of the magnitude of the resulting A–J inefficiency, Courville estimated the percentage deviation of actual cost from minimum cost at the actual output level for 105 of the 110 plants and found the average percentage deviation to be +11.4%, with a range of −0.6% to +40.6%. Using these figures, Courville estimated the total cost of A–J induced inefficiency to be approximately 437 million dollars in 1962.

Petersen (1975) based his study of regulatory effectiveness upon an econometric model which assumed cost minimization subject to the A–J rate-of-return constraint. Using this constrained cost minimization model, Petersen derived two testable implications of regulatory effectiveness, namely, that both total cost and the relative share of capital costs were decreasing functions of s, the allowed rate of return, assuming $s > r$, where r is the cost of capital to the firm. Under the assumption of an exogenous output quantity, Petersen derived a number of reduced-form specifications using a translog cost function and three separate measures of regulatory tightness: a dummy variable differentiating between states with and without statewide regulatory commissions, a dummy variable differentiating between state regulatory commissions using fair value versus original cost methods of rate base valuation, and lastly, a continuous measure of the difference between the allowed rate of return and the cost of capital based upon estimates of the return to equity capital. With respect to the first two measures, the dummy variables, it was assumed that utility regulation was more effective in states having regulatory commissions and in states using original cost valuation.

Petersen's sample consisted of data for 56 steam-generating plants which had experienced at least a 50 percent expansion during the period 1960 to 1965. Annual observations for the 3-year period 1966 to 1968 were used.[32]

In general, Petersen found that the derived implications of the constrained cost minimization model with respect to the influence of regulatory "tightness" upon both total costs and the relative share of capital costs were upheld by his statistical results. More specifically, the regulated state dummy variable was significant in both the total cost regression equation and the capital cost share equation, although the fair value dummy variable was not significant in either. The implication of this result is that state regulation does matter, i.e., is effective, although the method of rate base valuation does not appear to matter, at least with respect to its influence upon the level of costs. In the two regressions in which the third measure of regulatory "tightness", a measure of the difference between the allowed rate of return and the cost of equity capital, was used, Petersen found that reducing the allowed rate of return towards the firm's cost of capital resulted in significant increases in both total costs and the cost share of capital. The implication of this result is that rate-of-return regulation is effective and that changes in the allowed rate of return, s, have a significant impact upon the degree of cost inefficiency induced by the A–J regulatory constraint.

By way of comparisons among our study and the three studies described above, each of the four studies found evidence of a significant or binding A–J regulatory constraint so that rate-of-return regulation can generally be judged to be effective. The unanimity of these results is all the more impressive given the rather different approaches used, namely, both our study and Spann's tested for regulatory effectiveness by means of direct tests on λ, while Courville and Petersen based their tests on derived implications of the Averch–Johnson model. Nevertheless, there are also some differences among the four studies, differences which cast some doubt on the validity of the results of the three previous studies.

[32]Thus, the sample used suggests that a mixed *ex ante–ex post* technology was estimated since both new plants and old plants with significant expansion were included. Since total plant data was used to measure the input and cost variables, the result was probably a general confounding of both *ex ante* and *ex post* production technologies.

These differences entail specification problems with the econometric models used in all three studies, problems basically stemming from the specification of λ and the assumption, implicit or explicit, of exogeneity. These are also a number of minor problems.

The basic problem with the Spann model is its implicit assumption of a constant λ, constant, that is, across observations and, hence, across firms. As we have seen in this paper, the first-order conditions for constrained profit-maximization imply that λ, an index of the effectiveness or tightness of the regulatory constraint, is a function of input and output prices and, hence, in general will vary across firms. Thus, the proper econometric specification is one in which λ is treated as an endogenous variable and is allowed to vary across firms. In contrast, the assumption of a constant λ is likely to result in specification error and, hence, in a biased test of regulatory effectiveness. Spann clearly recognized this problem when he wrote:[33]

> There is a problem with the tests discussed here which does need to be mentioned. The term λ is treated as a parameter when it is really a variable within the model. The equations estimated are a subset of a more detailed simultaneous equation model which would include not only the share equations reported here, but the output demand equations and the production function as well. Within such a model λ could be treated as a variable, taking a different value for each observation in the sample. Unfortunately, λ is not directly observable. Thus, in the estimations that follow, the estimates of λ should be viewed as estimates of $\bar{\lambda}$, or the average of λ for the sample.

Unfortunately, this interpretation does nothing to remove the source of the bias so that his results may simply be biased estimates of $\bar{\lambda}$.[34] In addition, the fact that λ is unobservable does not present an insurmountable problem since the profit function approach used in our study allows λ to be treated as both endogenous and unobservable while still enabling the λ_i's to be estimated.

A second fundamental problem which is common to all three previous studies is the inadequate treatment of simultaneous equation issues. In both the Spann and Courville studies, input quantities were used as

[33]Spann (1974, p. 44).

[34]In a footnote, Spann states that regressions were run which yielded estimates of λ for each observation and that these results did not indicate very much variation in the estimated λ's. It appears, however, that λ was implicitly treated as a constant in the first part of these estimations so that misspecification bias is likely to have contaminated these results as well. At any rate, these results were not shown in the published results.

right-hand variables in the regression equations while Petersen used output quantity as well as input prices in his cost function specification. As we have seen, input quantities are endogenous variables to the regulated firm so that the failure of both Spann and Courville to estimate the full set of simultaneous equations probably led to biased estimates of the parameters in their regression equations. Since these estimates are crucial to their tests of regulatory effectiveness, the results of these tests may be suspect also. While Petersen's cost function specification represents an improvement over that of the other two studies, in that input prices rather than quantities appear on the right-hand side, the assumption of an exogenous output quantity may present similar kinds of simultaneous equation problems. Although Petersen offers several arguments in support of this assumption, the brief examination of the rate-setting procedure used by most regulatory commissions in this study supports the contention that output price, rather than output quantity, should be regarded as being exogenous to the regulated firm. If this is so, the profit function specification offers an improved specification to that of the cost function, unless the simultaneous equation problems are adequately dealt with in the latter.

In addition to these fundamental problems with respect to the proper specification of the Averch–Johnson model of the regulated firm, there are a number of other problems with these previous studies. In the Spann study, a constant cost of capital, equal to 0.056 based on the Litzenberger–Rao results for electric utilities, was assumed for all firms.[35] Thus, differences in the cost of capital across firms due to variations in risk, tax treatment and equipment prices was assumed away. Since the cost of capital is an extremely important variable to an electric utility, given the capital-intensive nature of the production technology and the nature of the regulatory process which focuses upon the rate-base, i.e., upon capital inputs, it would seem preferable to allow the cost of capital to vary across firms. A basic problem with the Courville study is that the effects of the regulatory process do not enter his regression equation (32) explicitly so that his estimates of the production technology of a regulated firm are likely to suffer from specification error in the presence of a binding regulatory constraint.[36] Since these estimates are used in his statistical test of regulatory effectiveness, his results may be questionable.

[35]From Litzenberger and Rao (1971).

[36]Courville's rejection of the null hypothesis of ineffective regulation reinforces this suspicion.

A minor problem with the Petersen study is that the econometric specification, using a translog cost function, is inconsistent with the theoretical model. Assuming that the translog cost function is interpreted as a second-order Taylor series approximation to some underlying, but unknown, cost function, the proper specification would be in terms of P_K (our r) and s, rather than P_K and $(s - P_K)$. Petersen's specification implicitly assumes that the coefficients associated with s and P_K in the expression $(s - P_K)$ are of equal absolute magnitude, a restriction which is not derivable from the Averch–Johnson model.[37] In addition, Petersen's econometric model could have been estimated more efficiently by including the cost function and two of the three input cost share equations in a simultaneous equation system. Although a revised version of Hotelling's lemma, pertaining to the regulated cost function instead of the profit function, would have to be used, the extra information introduced in the form of inter-equation restrictions would increase the efficiency of the estimation procedure.

In contrast with these three previous studies, the present study of regulatory effectiveness has used an econometric model explicitly derived from the Averch–Johnson model of a regulated firm. This model, derived from a regulated profit function which assumes the price of output, input prices and the allowed rate of return to be taken as given by the regulated firm, allows λ, the Lagrangian multiplier associated with the rate-of-return constraint, to be treated as an endogenous variable. In addition, the data used in this study covers a more extensive period of time. Thus, both the model and the data used in this study are improvements over those of previous studies and thus serve to significantly reinforce the general conclusion of all four studies, namely, the effectiveness of rate-of-return regulation in the electric utility industry.

6. Conclusions

Using a revised version of the Averch–Johnson model of the regulated firm in which both the output price and the allowed rate of return are assumed to be exogenous to the firm, this study has attempted several

[37]In addition, the s^2 term has been left out of the translog approximation. An alternative approach would have been to use the shadow price of capital for the regulated firm, our r'' in (10), since we have argued that the regulated firm can be regarded as if it were an unregulated firm facing input prices of w' and r''. Unfortunately, however, this causes λ, an endogenous variable, to appear on the right side of the cost function and thus raises simultaneous equation problems.

tests of regulatory effectiveness within an econometric framework which permits λ, the Lagrangian multiplier associated with the rate-of-return constraint, to be treated as an endogenous variable. Thus, this model allows for variations in regulatory effectiveness across both firms and time, a specification which is consistent with the theoretical implications of the regulated firm.

Our statistical tests indicate support for both the hypothesis concerning the general effectiveness of rate-of-return regulation and the hypothesis that such effectiveness appears to vary widely across different firms at the same point in time and across firms at different points in time. Since λ, a measure of regulatory effectiveness, is an endogenous variable which is itself a function of output and input prices, including the allowed rate of return, it is not surprising to find such variation. Thses results indicate that the efficiency implications of rate-of-return regulation may be significant, since the regulated firm does appear to respond to this regulation in a manner consistent with the general implications of the Averch–Johnson model, and that further research in this area should be based upon models which allow for variations in regulatory effectiveness. In addition to the results presented in our study we have also argued that our results are consistent with some findings presented recently by Joskow (1974) in an attempt to refute the A–J model. This would seem to further support the need for additional research since Joskow's data does not appear to be capable of differentiating between the two alternative models of the regulatory process, his institutional model and the A–J model.

Finally, it should be clear that further refinements in both the models and the econometric specifications used are needed. The model used in this study was a simple static equilibrium model which did not allow for either regulatory lag or for the possibility of interdependencies between the regulatory commission and the regulated firm. Since there is some indication that both of these effects may play roles in the actual process of rate-of-return regulation, future econometric research should focus upon including these effects within the models used. In addition, there is a need to extend this line of economic research into other areas of rate-of-return regulation as well. Up to the present time, all of the empirical research on the Averch–Johnson model has concentrated upon the electric utility industry, primarily because of the simplicity of the production process and the availability of data. We are in danger, however, of basing our knowledge of rate-of return regulation upon results of a single industry which, although an integral part of the overall economy, may not be typical of regulated industries in other respects. Future research, therefore, should concentrate upon expanding both the complexity of the model as well as the scope of empirical work. It would appear that the framework used in this study offers a fruitful approach for much of this future research agenda.

7. Appendix

7.1. *An Alternative Derivation of the Derivative Properties of the Regulated Profit Function*

The derivative properties of the regulated profit function (17)–(19) can also be derived using the comparative statics results in Table 1 as follows. We first note that the profit function for the regulated firm can be written as

$$\pi^*(w',r',s') = (s' - r')K^*(w',r',s'), \tag{35}$$

using the UOP profit function, by substituting (4) into (3) and using the optimal quantity of K, K^*. The regulated profit function (35) can now be differentiated with respect to each of the normalized arguments, w', r' and s', yielding

$$\begin{aligned} \partial\pi^*/\partial w' &= (s' - r')(\partial K^*/\partial w'), \\ \partial\pi^*/\partial r' &= (s' - r')(\partial K^*/\partial r') - K^*, \\ \partial\pi^*/\partial s' &= (s' - r')(\partial K^*/\partial s') + K^*. \end{aligned} \tag{36}$$

Expressions for $\mathrm{d}K^*/\mathrm{d}w'$, $\mathrm{d}K^*/\mathrm{d}r'$ and $\mathrm{d}K^*/\mathrm{d}s'$ are available from the comparative statics results summarized in Table 1. Substituting these results into (36) and noting from (9) that $(s' - r')/(s' - F_K) = 1 - \lambda$, we have

$$\begin{aligned} \partial\pi^*/\partial w' &\equiv \pi_{w'} = -(1 - \lambda^*)L^*, \\ \partial\pi^*/\partial r' &\equiv \pi_{r'} = -K^*, \\ \partial\pi^*/\partial s' &\equiv \pi_{s'} = \lambda^*K^*. \end{aligned} \tag{37}$$

which are precisely the results shown in (17)–(19).

7.2. *Additional Properties of the Regulated Profit Function*

Equations (37) can now be used, along with the results in Table 1, to derive additional properties of the regulated profit function, $\pi^*(w',r',s')$.

These results are summarized below, followed by the respective derivations:

(i) $\pi_{w'w'} \lesseqgtr 0$; $\pi_{r'r'} = 0$; $\pi_{s's'} < 0$,

(ii) $\pi_{w'r'} > 0$; $\pi_{w's'} \gtreqless 0$; $\pi_{r's'} > 0$,

(iii) The regulated profit function has ambiguous curvature properties in w' and s', w' and r', and r' and s'.

Differentiation of (37) with respect to each of w', r' and s' yields

$$\pi_{w'w'} = L(\partial\lambda/\partial w') - (1-\lambda)(\partial L/\partial w')$$
$$= \frac{(1-\lambda)[L^2\Delta - 2L(s'-F_K)F_{KL} - (s'-F_K)^2]}{(s'-F_K)^2 F_{LL}} \lesseqgtr 0, \quad (38)$$

$$\pi_{r'r'} = -(\partial K/\partial r') = 0, \quad (39)$$

$$\pi_{s's'} = K(\partial\lambda/\partial s') + \lambda(\partial K/\partial s')$$
$$= \frac{(1-\lambda)K^2\Delta - 2\lambda K(s'-F_K)F_{LL}}{(s'-F_K)^2 F_{LL}} < 0, \quad (40)$$

$$\pi_{w'r'} = -(\partial K/\partial w') = L/(s'-F_K) > 0, \quad (41)$$

$$\pi_{w's'} = K(\partial\lambda/\partial w') + \lambda(\partial K/\partial w')$$
$$= \frac{K(1-\lambda)(L\Delta - (s'-F_K)F_{KL}) - \lambda L(s'-F_K)F_{LL}}{(s'-F_K)^2 F_{LL}} \lesseqgtr 0, \quad (42)$$

$$\pi_{r's'} = K(\partial\lambda/\partial r') + \lambda(\partial K/\partial r') = K/(s'-F_K) > 0, \quad (43)$$

where $\Delta = F_{KK}F_{LL} - F_{KL}^2 > 0$, that is, assuming $F(K,L)$ is strictly concave which rules out the possibility of increasing returns to scale. This proves the properties summarized in (i) and (ii) above.

The proof of the curvature properties of the regulated profit function summarized in (iii) above follows from using the above results to calculate the signs of the appropriate Hessians as follows:

$$\pi_{r'r'}\pi_{s's'} - \pi_{r's'}^2 = -K^2/(s'-F_K)^2 < 0, \quad (44)$$

$$\pi_{w'w'}\pi_{r'r'} - \pi_{w'r'}^2 = -L^2/(s'-F_K)^2 < 0, \quad (45)$$

$$\pi_{w'w'}\pi_{s's'} - \pi_{w's'}^2 = \frac{2(1-\lambda)\lambda K(1+LF_{KL}) - (1-\lambda)^2K^2F_{KK} - \lambda^2L^2F_{LL}}{(s'-F_K)^2 F_{LL}}$$
$$\lesseqgtr 0. \quad (46)$$

Thus, the regulated profit function has ambiguous curvature properties in all three pairwise combinations of w', r' and s'.