Assessing Evidence for the Averch-Johnson-Wellisz Effect for Regulated Utilities

Stephen M. Law

Economics

Mount Allison University

This paper provides empirical support for the notorious observation of Paul Joskow (2005) that the exploration of the Averch-Johnson-Wellisz (AJW) effect over the past fifteen years has been a waste of time and effort. Entering the title of the paper by Averch & Johnson (1962) into the search engine for JSTOR yields 3853 items. Using a non-random selection of 130 peer-reviewed journal articles published since Averch & Johnson (1962) and Wellisz (1963), I find that almost 40% fail to provide the results of tests for the necessary pre-conditions before making assertions about the AJW effect or related economic results of regulation. Further, close examination of the empirical results provided in any remaining articles which claim to provide evidence of the AJW effect suggests that the AJW effect could not be present due to (1) single- rather than multi-period estimation, (2) incorrect capital price calculations, (3) problematic definitions of output, or (4) the reason that for the firms in these regulated industries, installed capital assets are complements to other inputs. Network industries, which typically have very high capital to labour ratios under any regulatory regime due to the nature of the industry, are frequently characterised by complementarity of capital and labour. Thus the kind of industry which is most often regulated may also be the kind of industry in which we are least likely to see the AJW effect. In general, there is little evidence to suggest that there was ever an AJW effect.
Introduction

The purpose of this paper is to present the theoretical and empirical support for the Averch-Johnson-Wellisz (AJW) effect in network utilities under rate-of-return (ROR) regulation, or, more accurately, to document the lack of evidence that has been used to support the common claim that the AJW effect has had widespread and important impact on utilities. The primary AJW proposition states that regulatory policy that requires the regulated firm to choose an output price to earn no more than an allowed rate of return, based on the level of installed capital, will create an incentive for regulated firms to choose a capital-labour ratio which is higher for the given output level than would be chosen without the regulation, that is, that ROR regulation causes distortions in input levels.

A necessary condition for the AJW effect is that capital and labour are to some significant extent substitutable in the production process. Otherwise, if they are complements, then the impact of the AJW effect is smaller and if the two inputs are perfect complements then there is no scope for the AJW effect whatsoever. Another necessary condition is that the regulatory constraint is binding on the operations of the firm which may not be the case if, in the wider set of regulatory interactions that occur before and/or after the setting of the allowable rate of return, there is some channel through which the firm can influence the outcome.\(^1\) In addition, for the AJW effect to matter, the regulator must not have taken steps, by adjusting the time lag between regulatory hearings for example, to mitigate the effect.\(^2\) Many theorists and empirical researchers believe that the AJW

---

\(^1\) These necessary conditions have also been noted by other commentators, for example by Kolpin (2001) in the *Review of Industrial Organization*:

If the monopolist perceives there is even an implicit policy in which past behaviour may influence future allowed returns, the AJ effect does not apply and one need not expect cost inefficiency to be observed. Another class of examples emerges when production and/or profits fail differentiability. For instance, it is easy to construct examples in which a firm endowed with Leontief production technology will continue to employ labor and capital in efficient proportions when exposed to rate-of-return regulation. More generally, any scenario in which the marginal productivity of “capital” varies discontinuously with the availability of those inputs necessary for their operation is subject to the failure of the AJ effect. Kolpin (2001), page 181.

\(^2\) See Church & Ware (2000), page 849.
effect, even where it exists, is likely to be of very small impact.\textsuperscript{3} And finally, an examination of the results of tests of complementarity and separation in the literature which reports empirical examinations of the production functions of regulated firms suggests that network industries, which have very high capital to labour ratios under any regulatory regime due to the nature of the industry, might be generally characterised by complementarity (or at least a low degree of substitution) of capital and labour. The kind of industry which is most often regulated is also the kind of industry in which we are least likely to see the AJW effect. And so, while there may be other reasons to “not prefer” rate-of-return regulation, the AJW effect is not one of them.

Although the theoretical structure developed by Averch & Johnson (1962) and Wellisz (1963) has been carefully disseminated in courses in Industrial Organization and Regulatory Economics ever since, the necessary empirical conditions for the AJW effect have been generally ignored and typically remain untested in articles that nonetheless claim its pertinence.\textsuperscript{4}

\textsuperscript{3} See, \textit{inter alia}, Baumol & Klevorick (1970).

\textsuperscript{4} I am not immune to this accusation (\textit{mea culpa}). In 1999, I published an article in the \textit{Canadian Journal of Economics} in which I wrote that

The awareness on the part of regulated firms of the propensity of a regulator to appropriate some of the gains from more cost-efficient operations results in a reduced incentive to acquire cost-reducing equipment (see Besanko and Spulber 1992; Spiegel and Spulber 1994). This incentive effect, the Besanko-Spulber effect, stands in contrast to the Averch-Johnson effect, which typically arises under rate-of-return regulation, which the CRTC has not instituted, or capital expenditure allowances, which it has. The Averch-Johnson model “examines how a regulated firm picks its inputs when the regulator exerts no control over this choice and the firm is permitted a rate of return on capital exceeding the cost of capital ... [thus, since] capital investment expands the rate base on which the firm is allowed an excess rate of return ... this induces the firm to select excessive capital-labour ratios” (Laffont and Tirole 1993, 33). The structure of the CRTC’s rules is biased toward capital investment: the price cap reduces the possibility that cost-savings from new equipment will be eroded by corresponding rate decreases; the CAPEX clause directly rewards investment with rate increases; and the rate increases permitted when the rate-of-return is below the benchmark provide an additional incentive for CATV firms to hold a larger stock of assets.

Despite presenting results from estimation over data that \textit{might} have provided some indication of whether capital and labour were substitutes, the article is silent on this matter. Later work in this area (Haghiri, Law & Nolan (2004)) suggests that capital and labour are additively separable. This implies (weakly) that labour and capital are likely to be complementary inputs rather than substitutes.
In 1973, almost eleven years – and many published studies – after the publication of the original 1962 article, Leland Johnson observed that the question remains about the importance of overcapitalization and cross-subsidization in reality. Are the Averch-Johnson effects merely an intellectual curiosity, or do they describe serious distortions in the behavior of regulated firms? Unfortunately the answer is not clear. It is not enough to compare the behavior of regulated and unregulated firms... Johnson (1973), page 91.

By 2005, Joskow’s answer to Johnson’s question appears to be that the AJW effects are an intellectual curiosity and this response is confirmed here.

One purpose of this paper is to document the prevalence of this failure to provide evidence to support an assertion of the importance of the AJW effect. Another is to connect observations about regulatory theory to this discussion. And, finally, this paper provides an assessment of Joskow’s claim that wasted were the efforts over the past fifteen to come to terms with the work of Averch, Johnson, and Wellisz. It is worthwhile to begin with the original statement of Averch and Johnson and follow its evolution over the subsequent decades.

2. The Averch-Johnson-Wellisz Model and Interpretations

2.1 The AJ Model: Averch & Johnson (1962)

In 1962, Harvey Averch and Leland Johnson published a paper in the American Economic Review which set out a model for a regulated monopolist producing an output, \( z = z(x_1, x_2) \), inverse demand function given by \( p = p(z) \), and two inputs, physical capital, \( x_1 \geq 0 \), and labour, \( x_2 \geq 0 \), such that \( z(x_1, 0) = z(0, x_2) = 0 \), \( \frac{\partial z}{\partial x_1} > 0 \), and \( \frac{\partial z}{\partial x_2} > 0 \). With factor prices, \( r_1 \) and \( r_2 \), profit is given by

\[
\pi = pz - r_1 x_1 - r_2 x_2.
\]

Given an acquisition cost for capital of \( c_1 \), current depreciation, \( u_1 \), and cumulative depreciation, \( U_1 \), the constraint for rate-of-return regulation can be written as
where $s_1$ is the maximum allowable rate of return. Averch & Johnson set $U_i = u_i = 0$ and $c_i = 1$ for convenience and note that if $r_1 > s_1$ the firm exits. Assuming $r_1 < s_1$ they define

$$L(x_1, x_2, \lambda) = p_1 x_1 - r_1 x_1 - r_2 x_2 - \lambda[p_2 - s_1 x_1 - r_2 x_2]$$

and note that if $\lambda = 0$, the firm is not constrained by the regulation, if $\lambda = 1$, $r_1 = s_1$, the firm is so constrained that any combination of $x_1$ and $x_2$ such that (1) holds is a solution and, finally, if $0 < \lambda < 1$, then the firm is constrained by the regulation and “the input of $x_1$ is such that...its use is expanded beyond the point at which its marginal cost would be equal to its marginal value product”.^5 They also note that “the extent to which the regulation affects output depends on the nature of the production function [and] if it involves fixed proportions, i.e., $\min\left\{\frac{x_1}{a}, \frac{x_2}{b}\right\}$, the regulated firm is constrained to the efficient expansion path.”^6

### 2.2 Geometric Interpretation: Zajac (1970)

In 1970, Zajac published an article in the *American Economic Review* presenting a geometric interpretation of the AJW model. The motivation for the article, according to Zajac, was that “unfortunately, Averch and Johnson carried out their analysis using the rather abstract tools of nonlinear programming and the Kuhn-Tucker theorem [and hence] its detailed understanding has thereby been denied a large number of persons concerned with regulation for whom these tools are foreign and uncomfortable”.^7 Zajac recasts the regulatory constraint for a firm producing output, $q$, as:

---

^5 Averch & Johnson (1962), page 1056.

^6 Corey (1971), at page 364, contains a diagram and a discussion of the absence of the AJW effect when inputs are complements.

^7 Zajac (1970), page 117.
where \( x_1 = K \) is capital with average factor cost of \( r_1 = i \), \( x_2 = L \) is labour with average factor cost of \( r_2 = w \) and \( f \) is the “fair rate of return” imposed as a maximum by the regulator. Since profit is thus \( \pi = pq - iK - wL \) the regulatory constraint is \( \pi \leq (f-i)K \). Zajac then demonstrates that a profit-maximizing firm will choose the largest level of capital, \( K_{\text{MAX}} \), that satisfies the regulatory constraint. Over the following four decades, instructors of courses in Industrial Organization have either labored to reproduce the Zajac diagrams on blackboards, photocopied the figures for class hand-outs for their students, or required students to seek out Zajac’s article on their own. The diagrams are provided in Appendix 1 to this paper.

Zajac notes that a key assumption is that the rate of return set by the regulator exceeds the cost of capital (i.e., \( f > i \)) but does not make any comment about the underlying technology of production. Zajac demonstrates that the firm does not have an incentive to acquire useless capital. An implication of this result is that if the underlying technology involves fixed proportions, the constraint curve depicted in Figures 2 and 3 will lie over the set of efficient points and the \( K_{\text{MAX}} \) point will be on the expansion path of efficient points, that is, there will be no Averch-Johnson Effect. This result was noted by Averch and Johnson in their original article but, importantly, was omitted from Zajac’s more accessible treatment of the AJW model. Importantly, because it was generally to Zajac (1970) that scholars and students of regulation turned when seeking “a detailed understanding” of the analysis of Averch and Johnson.

2.3 Restatement and Correction: Baumol & Klevorick (1970)

The other article which was essential for understanding the AJ Effect was also published in 1970. Baumol & Klevorick’s article in the *Bell Journal of Economics and Management Science* begins by noting that the model in Averch & Johnson (1962) is similar to one published at almost
the same time by Wellisz in the *Journal of Political Economy*. Baumol & Klevorick suggest that “the phenomenon that emerges from the A-J theorem may not be of very great significance in practice” and “it is at least plausible that other potential sources of difficulty in the regulatory process dwarf the consequences of the distortion in the capital-labour ratio that the model predicts.” Baumol & Klevorick restate the regulatory constraint as:

$$ s = r_1 + v \geq \frac{pZ - r_2 x_2}{x_1} $$

with $v \geq 0$, assuming with Averch & Johnson, $U_1 = u_1 = 0$ and $c_1 = 1$, and also making the standard assumption that “$r_1 + v$ is taken to be less than the [unregulated] profit-maximizing rate of return”.

Baumol & Klevorick extend the logical structure created by Averch and Johnson and conclude that the original authors and those who took it up over the subsequent years (between 1962 and 1970) either implied or assumed the following propositions:

**Proposition 1.** The profit-maximizing firm under regulatory constraint will tend to use a capital-labor ratio different than that which minimizes cost for its output level.

---

8 Unlike the addition of “Boiteux” to “Ramsey-Boiteux” – in Caillaud, Guesnerie, Rey & Tirole (1988) and in *A Theory of Incentives and Procurement* by Jean-Jacques Laffont and Jean Tirole (1993) – which has been largely successful, as measured by increased citations of the work of Boiteux, the observation of William Baumol and Alvin Klevorick does not seem to have convinced many subsequent authors to adopt as the name of the regulatory model they discuss the Averch-Johnson-Wellisz Model or its main result the AJW Effect. In the years immediately following the publication of the AJW papers, some scholars were careful to mention both – examples include Westfield (1965) and Shepherd (1966) – some authors were careful to cite both, even if the contribution of Wellisz was not noted explicitly in the text – examples include Takayama (1969) and Corey (1971) – but despite a few rare exceptions – such as Filer & Hollas (1983) and Evans & Garber (1988) – references to Wellisz drop off sharply after 1971 even while research into the “AJ effect” continues to (at least) 2008. In notable contrast to this observation are only two articles – the first in the *American Economic Review* by Callen (1978) and the second in *Land Economics* by Berg & Tschirhart (1995) – which refer to the AJW Model.

9 Baumol & Klevorick (1970), page 164, emphasis added.

10 Baumol & Klevorick (1970), pages 165, 166, 168, 175, and 180.
(Alleged) Proposition 2. *The capital-labor ratio of the regulated firm will be larger than that of the unconstrained profit-maximizing monopolist.*

Proposition 3. *The capital-labor ratio of the regulated firm will be larger than the one that minimizes costs for the output level that it elects to produce.*

(Alleged) Proposition 4. *The regulated firm will produce an output larger than that which maximizes profits.*

Proposition 5. *For the firm that seeks to maximize total profit subject to the regulatory constraint, we have* $dx_1 / ds < 0$ for $r_1 < s < r_m$ [where $r_m$ is the rate of return obtained at the (unconstrained) profit-maximizing input-output combination]. In other words, the greater the difference between $s$ the regulatory fair rate of return and $r_1$ the cost of capital (since we are increasing $s$ holding $r_1$ constant), the smaller will be the value of $x_1$, the firm’s use of capital.

Proposition 6: *The sales-maximizing firm under rate-of-return regulation is motivated to use a labor-capital ratio greater than that which minimizes cost for the output level it chooses to produce.*

Baumol and Klevorick provide restatements of the proofs of Propositions 1 and 3. As might be supposed by the use of “alleged” they show that Propositions 2 and 4 are false and these points (1, “true”; 2, “false; 3, “true”; 4, “false”) have been incorporated into standard regulatory theory and doubtless have provided the basis for a large number of exam questions in undergraduate courses in Industrial Organization or the Economics of Regulation.

In demonstrating Proposition 3 it would appear at first that Baumol and Klevorick have failed to note the impact of having perfect complementarity between inputs, noting only that for the
regulated firm “$z_1 / z_2$, the marginal rate of substitution of capital for labor, is now below the ratio of input prices $r_1 / r_2$ [and] with diminishing marginal rate of substitution of capital for labor for a fixed output level, this can occur only as the result of relative increase in the use of capital.”

However, later in the paper, we find their statement that

Depending on the marginal rate of substitution between $x_1$ and $x_2$, the employment of labor may either increase or decrease as $s$ gets closer to $r_1$. If labor and capital are complementary in the gross revenue function, then as the quantity of capital used by the firm increases its use of labor will also rise. But if capital is a substitute for labor in producing revenue, then $x_1$ and $x_2$ will move in opposite directions in the A-J model.\(^{12}\)

and also that “since capital and labor can either be complements or substitutes in producing output....we simply cannot conclude that the A-J proposition 4 is always valid.”

In their discussion of an article by Bailey & Coleman which develops a model of regulatory lag, Baumol and Kleverick connect the timing of regulatory reviews to the regulated firm’s choice of capital level: “in practice, the rate-of-return constraint is not enforced continually” and “one might surmise that ... [regulatory lag] serves to weaken the A-J input-proportion effect, at least to some extent.”\(^{14}\) This observation can be made more powerfully after considering the later contributions of Joskow, Spulber, Laffont, Tirole, and others who have sought to embed the regulatory constraint within a more general model of regulator-firm interactions and we return to this topic below. Nonetheless, in 1970, Baumol and Kleverick concluded that the conventional treatment of effect of the regulatory constraint – without considering the regulatory lags or specifying more carefully the incentives of the firm – led to an overemphasis on the AJW effect as a source of inefficiency. They note that:

\(^{11}\) Baumol & Kleverick (1970), pages 167-168.

\(^{12}\) Baumol & Kleverick (1970), page 177.

\(^{13}\) Baumol & Kleverick (1970), page 178.

\(^{14}\) Baumol & Kleverick (1970), page 182.
“the A-J overcapitalization is an example of the inefficiencies emphasized in the more conventional analyses. But even if it occurs in practice it does not seem likely to produce effects that are very serious.”

On the other hand, researchers do not shy from mentioning AJW Effect when mentioning regulation; see Mayo and Flynn (1988) page 322.

“The point is simply that while regulation may well be suspected of being the source of some non-negligible inefficiencies in the economy, it is not clear that the phenomenon encompassed by A-J analysis is the most disquieting of these.”

And further research has supported this point, as we shall see.

2.4 Modern restatements: examples Caputo & Partovi (2002), Church & Ware (2000)

Caputo & Partovi (2002) provide a concise derivation of what Baumol & Klevorick (1970) labelled Proposition 5 and show the equivalence of a number of different conditions, with economic content, that may be imposed to achieve this proposition. In so doing, they closed the debate between Akira Takayama and Mohamed El-Hodiri on the one hand and Israel Pressman and Arthur Carol on the other regarding the (theoretical) existence of the effect presented in Proposition 5. This effect, though, is not what is commonly taken to be the A-J effect.

For an excellent presentation of the AJW model, interested readers are encouraged to turn to Church & Ware (2000) for their discussion of cost-of-service regulation, pages 841-852. Church and Ware present the AJW model and then go on to consider three mitigating factors: regulatory lag, prudence and facilities reviews, and regulatory risk. Under this last topic, they note that

---

15 Baumol and Klevorick (1970), page 188.
16 Baumol and Klevorick (1970), page 189.
17 Takayama (1969), El-Hodiri & Takayama (1973)
18 Pressman & Carol (1971), Pressman & Carol (1973)
The A-J model assumes that capital expenditures are not sunk. In fact utilities are regulated typically because they are natural monopolies due to large sunk capital investments. Church and Ware (2000), page 850.

This point forms the basis for some of the observations provided below, drawn from empirical studies of regulated utilities.

2.5 Theoretical Problems with the AJW Propositions

The first set of problems with the Averch-Johnson-Wellisz model were in its construction. Many of these problems were corrected in subsequent reformulations such as Zajac (1970), Stonebender (1972) and Baumol & Klevorick (1970), who more carefully constructed the analysis, and even more recent papers such as Borrmann and Finsinger (2006) who calculate the range for the multiplier in the AJW model assuming that the profit function is single-peaked. Some authors maintained that the AJW proposition is fundamentally flawed, for example:

It has been shown by L. Courville that a proof of overcapitalization requires the additional assumption of strictly convex isoquants....Thus, Courville has strengthened our contention “that the very assumptions used to prove the A-J effect . . . require an assumption that the A-J effect exists in the first place” [Quotation drawn by Pressman & Carol from p.210 of Courville’s Carnegie-Mellon PhD Dissertation] Pressman & Carol (1973), page 238.

However, the requirement of strictly convex isoquants has not been found to be overly restrictive and, indeed, most researchers make this assumption, implicitly or explicitly.

A second set of problems to arise with the AJW Model is more serious. The fundamental prediction of the model is that in the presence of rate-of-return regulation “given $Q, K/L$ is too high”, that is, the capital-labour ratio, $K/L$, chosen by the firm will exceed the ratio that would be chosen by an unregulated firm, were that firm to be producing the same level of output, $Q$ (which it likely would not). But what if there is rate-of-return regulation and the most appropriate model of firm-regulator interaction predicts something else? Besanko & Spulber (1992) construct a model in which the firm chooses a level of capital and the regulator subsequently sets the price cap or allowed rate-of-return. Laffont & Tirole (1993) observe that price caps and rate-of-return regulation are basically
equivalent. Besanko & Spulber (B&S) predict that in the face of regulation, the firm chooses a lower than optimal level of cost-reducing capital because the regulator, acting second, would appropriate the returns from the capital.\(^{19}\) We are left with the following problem: the AJW proposition states that if there is rate-of-return regulation, given \(Q\), \(K/L\) will be “too high”; while in contrast, the B&S model suggests that if there is binding regulation, \(K/L\) will be “too low”, given \(Q\). The primary distinction is whether the firm moves first or the regulator moves first. This distinction is easy to make in a theoretical model but much more difficult to test empirically. In practice, both the regulator and the firm continually emit signals about their intentions and their assessments of market and cost conditions. Regulatory hearings provide a formal forum for the dissemination of these signals and there are additional dissemination channels such as press releases, announcements, and interviews. Typically, firms can anticipate regulatory changes before they are passed by formal procedures and regulators follow developments in the industry at trade shows and through business reports and reports to shareholders. The B&S model and other similar models which set the regulator and the firm or firms inside a regulation game form a more general framework for the evaluation of regulation and, more importantly for this paper, for the construction of empirically testable predictions about the behavior of firms. This wider set of possibilities does not rule out the AJW effect entirely since a richer model may deliver the AJW effect as a sub-case, one of many possible outcomes, but it reduces its general applicability.\(^{20}\) For this wider model to be useful, we need a way of identifying the essential timing characteristic which would give us the AJW result

\(^{19}\) See also Spiegel & Spulber (1994).

\(^{20}\) Liston (1993) discusses some examples of regulatory games that yield the AJW effect as a subcase:
Besanko (1984) argues that in a symmetric information environment the A-J model is unsatisfactory because it implicitly assumes that the regulator acts myopically. However, if the firm has private information about its productivity parameter, he suggests that the input bias could arise endogenously as part of the optimal regulatory regime. He constructs an adverse selection model in which the regulator uses a "graduated ROR" that decreases with the capital employed, as well as its power to monitor capital investments. Because less productive firms benefit more from increases in capital than more productive ones, the graduated ROR induces them to self-select. The resulting over-capitalization (i.e., the input bias) is viewed as an unavoidable, although welfare improving, consequence of lower information rents. (See also Baron’s (1990, 1380-81) presentation of Besanko’s model.) Liston (1993), Note 13, page 41.
versus the B&S result. Assuming we can distinguish the temporal sequence of interactions we could offer the following predictions, in the presence of ROR, if the regulator “moves first”, for a given $Q$, there is a higher $K/L$; if the regulator “moves second”, there is a lower $K/L$, given $Q$. More importantly, regardless of whether this temporal distinction can be made operational, while there may be reasons to suggest that regulators might want to steer clear of rate-of-return regulation, the automatic presumption that we get sub-optimal capital-labour ratios should be discarded.

Instead of recasting the AJW effect as a single outcome, a subcase, of a larger regulatory game, Joskow makes an observation similar to the one made here and provides a different solution to the problem, discarding the AJW model altogether:

“The Averch-Johnson model and its progeny have been replaced with a richer set of models of regulation, both normative and positive, that consider asymmetric information, political economy considerations, legal constraints on agency behavior and their effects on the incentive properties of regulatory mechanisms and ultimately on the behavior and performance of regulated firms. I have in mind here in particular work by Laffont and Tirole (1986, 1993) Baron and Besanko (1984), [Lewis and] Sappington (1988) and many others.” Joskow (2005), p.188.

A final complication, raised by theoretical analysis but most pertinent to empirical studies, is that even if the effect is found its source is not necessarily the AJW model.

A conclusion of the study is that the existing regulatory regime, which has as its primary instrument the determination of allowed rate of return on invested capital, provides utilities with incentives to invest in base-load capacity at levels that exceed the socially optimal level. Although this conclusion is similar to that derived in the seminal paper by Averch and Johnson (1962)..., it is derived from a model which is substantially different from theirs. Furthermore, our conclusions are stronger. Gal-Or & Spiro (1992), page 264.

We turn now to additional problems that arise when attempting to provide evidence of the AJW effect in empirical studies of regulated network utilities.

There is a fundamental problem in estimation. In the presence of regulation, a researcher may take published regulatory practices or posted rules and construct a regulatory constraint. The presence of the AJW effect is then assumed and the AJW model is used to obtain estimates of characteristics (such as input complementarity) of the regulated firm’s technology. This method is used fairly commonly in the literature. An example of a theoretical presentation of this approach can be found in Färe and Logan (1983). Färe and Logan observe that if we can write down the regulatory constraint that gives rise to the AJW model we can use the specific prediction of the AJW model to back out the regulated firm’s technology. They assume a particular response to the regulation and assume that the AJW effect is as predicted. Subsequent empirical use of this approach is based on the following binary test: either the regulation is binding and yields the AJW effect or it is not binding and there is no distortionary effect. If the regulation is not binding then results assuming that it is and results ignoring the regulation should not be statistically different. If the results are statistically different, then most researchers conclude that the regulation is binding and has an impact as predicted by Averch, Johnson and Wellisz. Among other problems, if the input distortion arises from some other source (e.g., cost pass-through rules) then the researcher may be left unable to perform the calculation required to back out the unregulated firm’s technology. Further, as Färe and Logan themselves note, “to reconstruct the rate-of-return regulated production function, it is necessary to have knowledge about the rate-of-return constraint as well as to know the rate-of-return regulated cost function.” So unless the assumed regulatory constraint captures exactly the interaction between the firm and the regulator, when we use this approach all conclusions are conditional on the accuracy of the formulation of the regulatory constraint. Given a sufficiently rich data set, we are able to examine past behaviour of regulated firms with the goal of determining the impact of regulation or uncovering features of the underlying technology. But there is a danger

---

21 This procedure sometimes involves estimating the Lagrangian multiplier ($\lambda$, from Equation 2 above) and testing whether it is significantly different from zero or from one.

within these projects of assuming the thing we have come to test (e.g., with a naïve application of the Färe and Logan methodology).

Beyond explanation, we wish to predict even if this exercise is complicated by the fact that our observations are coloured by the regulatory framework that is part of the specific data generating process that a particular study might be investigating. If we wish to predict, we must have accounted for the effect of the regulatory game. Does the firm move first? Or does the regulator? Do their responses change over decades as economists publish papers suggesting that their responses are driven by expectations of changes to rates-of-return, capital-labour ratios, or other important economic variables? We must be sure our predictions are robust to the choice of game.

Previous papers have attempted identification through the use of (1) time-series (before-and-after comparison); (2) cross-section (regulated-and-not comparison) plus combinations (panel data) and (3) less extreme versions of each of these involving variation in the intensity of regulation. Examples of these approaches (from Joskow & Rose (1989)) include Spann 1974 and Nelson & Wohar (1983).

But if identification is not so simple, if the data set employed does not deliver clean counterfactual scenarios for testing, then there is a problem with the assertion that rate-of return regulation has caused input distortion since the AJW model may not be responsible for the apparent result. There are many other possible explanations for “overcapitalization”, some of them more readily identifiable, econometrically, than others. An expectation of rising demand, for example. The testing of the AJW effect occurred during a specific period of history, during which many large corporations in the private sector that were unregulated also operated with high capital-labor ratios or used excess labor. Part of the reason for these features of economic organization in this time included: the rising power of labor so that these measures would be adopted to avoid strikes or to avoid unionization; management practices of the time; measures to reduce labor turnover where
workers had (unmeasured) human capital; and some part of an excess use of capital inputs may be the implicit creation of an efficiency wage by reducing the effort-level of an employee.

Joskow and Rose, commenting on the importance of identification when studying the impact of regulation, note that

Interactions of regulation with changing economic conditions may, when properly modeled, provide an additional way of identifying regulatory effects [Joskow (1974), Carron and MacAvoy (1981), Hendricks (1975), Burness, Montgomery & Quirk (1980), Greene & Smiley (1984)]. In particular, certain regulatory constraints may be binding under one set of economic conditions, but not under another. Implementing this approach requires particular attention to the nature of the regulatory process under study and how it works when economic conditions change. Joskow’s (1974) model of state public utility commission behavior provides an example of this approach. Joskow & Rose (1989), p. 1461.

Note that “paying attention to the nature of the regulatory process” differs materially from “assuming the nature of the regulatory model” which is the incorrect but more common approach. Joslow and Rose also claim that “estimates of firms’ production functions, combined with information on input prices, can be used to test whether regulated firms make cost-minimizing output choices”.

But where are the production functions obtained? Typically – and problematically – from estimates based on the data drawn from the actual experience of the regulated firm or firms. Joskow and Rose go on to note that although:

“estimating demand functions for regulated firms should present no particular difficulties....we are not as sanguine about cost or production function estimation. Estimates of production or cost functions from observed combinations of outputs, inputs, input prices, and costs tend to rely on a number of implicit assumptions, including equilibrium conditions and exogenous factor prices. These may be implausible for many regulated markets.” Joskow & Rose (1989), page 1463.

and, further,

“To the extent that one treats factor prices as exogenous [when they are endogenous], or fails to model explicitly direct regulatory constraints on production decisions, the resulting cost estimates can be quite misleading.” Joskow & Rose (1989), page 1463.
And what if the regulatory constraints are endogenous, the result of the game played by the firm and the regulator? Explicitly modelling the constraints is of not much use if these constraints are determined in a wider game. Using the AJW constraints to estimate features of the underlying technology is not a very viable approach if the underlying technology does not support the AJW model, if capital and labour are complements for example. We turn now to some of the empirical studies which grappled more or less carefully with these problems.

4. Representations and Assessments of the Averch-Johnson-Wellisz Model

“In my view, students of regulation of legal monopolies wasted at least 15 years extending the Averch-Johnson model of regulatory behaviour and trying to test it empirically without much success.” Joskow (2005), page 188.

Entering the title of the paper by Averch & Johnson (1962) into the search engine for JSTOR yields 3853 items (as of the end of May 2008). Once items which have no direct connection to the AJW model are removed by searching within the 3853 items for those with the authors’ names, this number drops to 425. Finally, selecting from the remainder those articles which seem to pertain more-or-less directly to the issues raised here (and adding some relevant studies found through other searches) yields a non-random sample of 130 journal articles, book chapters, or working papers. This sample forms the basis for the analysis and comments of this section of the discussion. It is intended to be at least somewhat representative of the response of economists to the AJW model. The sample itself is reproduced in Appendix 2 below.

4.1. “The Problematic Studies: Inconclusive or Incomplete”

Following in the tradition of McKay (1976) who assesses three empirical investigations of the AJW effect – and extending evaluative principles to theoretical papers – the items in the sample are examined to determine what they contribute to the debate on the existence and/or importance of the AJW effect. We consider three possible determinations in decreasing order of importance:
“problematic”, “not careful”, and “not thorough”. Using a non-random sample of peer-reviewed articles, we calculate a number of evocative statistics.

The least serious category is the “not thorough” grouping which contains all papers which do not explicitly present all of the cases and required conditions for the AJW effect. Of the 130 items considered here, 115 or 88.46% fall into this group.

We find we place an unfortunately large number of studies – well over one third – into the “not careful” category: 50 items or 38.46%. Theoretical papers were placed into this category primarily for being misleading or incomplete, especially in regard to implicit assumptions. The most common reason for a theoretical work to be considered “not careful” is if the substitutability of inputs was assumed without explicit mention of its importance. Empirical studies were placed in this category if the AJW effect is discussed (but not always asserted as “found”) without a presentation of any tests for necessary pre-conditions, especially tests on the complementarity of inputs.

If we consider as “problematic” items which are flawed or inconclusive, 7 of the 130 or 5.38% fall into this category. These are a subset of the “not careful” group discussed above. The primary reason for a paper to be placed in this category is for the AJW effect to be reported as “found” or “proven” either on the basis of an empirical study which is subsequently shown to suffer from methodological or data problems or for the AJW effect to be reported without any test for the presence of one or more of the economic conditions necessary for its existence. These papers, then,

---

23 Caputo & Partovi (2002) provide a precise, and convenient, set of statements for the economic conditions which may be assumed in order to establish the AJW model.

24 Rothwell & Eastman (1987) emphasize the importance of measuring the cost of capital carefully before drawing any conclusion about the AJW effect, noting that:

Different measures of the cost of capital yield different conclusions about the appropriateness of using models assuming the same financial and regulatory constraints as in Averch and Johnson [1962]. The realized rate of return was greater than the AFUDC rate for all years except 1981. However, when we use a market price of the cost of capital, the realized rate of return was less than the cost of capital.
go further than the “not careful” papers in that their authors claim to have provided evidence of the existence of the AJW effect, rather than simply discussing the AJW effect in passing. Given the widespread acceptance of the AJW effect, the test for inclusion in this category is biased: no empirical paper which fails to find the AJW effect is labelled problematic. Readers should adjust their posterior beliefs accordingly.

Another way for a paper to be considered “problematic” is if it presents one of the common misconceived versions of the AJW effect, that is, that the AJW propositions imply that the regulated firm will choose a capital stock that is too big, the regulated firm will pay a price for its capital that is too high, or that the regulated firm will acquire unproductive capital. None of these is correct.

Examples of problematic papers include Courville (1974), Spann (1974) and Petersen (1975) all of which are dealt with in McKay (1976) whose “single most important objection to these studies is that they neglect to take into account one of the basic assumptions which is made when production or cost functions are used to represent technological possibilities.” Since regulatory authorities require utilities to satisfy demand, “the use of annual energy as the output and either total plant cost or capacity as the measure of capital contradicts....the assumption that engineering suboptimizations have taken place so that the function gives the maximum output attainable with the given inputs.”

Gollop and Karlson (1980) also provide evidence to refute the finding of the AJW effect in these three papers. Gollop and Karlson

develop and apply a multiperiod econometric model....The resulting empirical description of the industry's technology and the estimated residential demand elasticity are consistent with the findings of other applied research. The important

from 1979 to 1982. This finding violates the financial constraint of the single-period Averch-Johnson model, suggesting that the model is an inappropriate description of electric utility behavior in the late 1970s and early 1980s. (This does not imply that the model is necessarily appropriate before the late 1970s.) Also, we found that the allowed rate of return was greater than the realized rate after 1976. Using the realized rate as a proxy for the allowed rate in single-period models will bias estimates of overcapitalization toward accepting the Averch-Johnson thesis. Rothwell & Eastman (1987) page 108.

25 McKay (1976), page 2.
empirical conclusion, however, is that we find no evidence of input distortion. Gollop and Karlson (1980), page 313.

When Gollop and Karlson restrict their model to a single period, the result confirms the Averch-Johnson hypothesis and supports the earlier research by Spann (1974), Courville (1974), and Petersen (1975). The results of the more general inter-temporal model, however, suggest that the above inference is the result of specification bias and not regulatory bias. This contrasts with the single-period model inference that inefficient producer behavior can be attributed to rate-of-return regulation. This "inefficiency," however, most likely is due to specification bias, not regulatory bias. Three of the four estimating equations in the single period and multiperiod models are identical. Only the optimizing condition with respect to capital is specified differently. Of course, it is precisely this first-order condition that is central to the evaluation of the Averch-Johnson hypothesis. Gollop and Karlson (1980), page 313.

Another, more recent example, is found in Saal & Parker (2001) who report finding an AJW effect in regulated water utilities in England and Wales.

Capital for labor substitution has been occurring during the 1990s, something consistent with the argument that, where economic regulation allows for a rate of return on investment at or above the cost of capital, incentives exist to overinvest (Averch and Johnson 1962). Regulation of the water industry in the 1990s seems to have failed to counteract this tendency. Saal & Parker (2001), page 87.

But the study does not present any tests for necessary conditions; in particular, the substitutability of labour for capital in the operations of water and sewer services is simply assumed in the analysis.26

4.2 Evidence for the AJW effect

Some papers provide a careful counter-factual and do not assume the AJW effect before attempting to find it or discover its effect on the production or cost functions. Of the many papers that claim to have found evidence of the AJW effect, only a few use methodologies that are sufficiently thorough or careful. Even these studies, taken together, are not conclusive. For

26 Although many of the utilities in their sample reduce labor input and increase capital input while increasing output, the authors provide no elasticities or statistics to support a conclusion of substitution.

Granderson & Lovell (1998) find the AJW effect in US Natural Gas Pipelines for 1977-1987 and, using much the same data set, Granderson & Linvil (1996) find evidence of the AJW effect for US Natural Gas Pipelines for 1981-1987 but not 1977-1980. The finding of the AJW effect is for the period just prior to the deregulation of this industry. It is possible that, at least toward the end of the sample period, as firms anticipated deregulation, more complex firm-regulator interactions were occurring than those contemplated by the AJW model.

A very persuasive finding of input bias in a regulated industry is provided by Atkinson, Färe & Primont (2002) for US Railroads, over the 1951-1975 period, but no claim that this input bias is the result of the AJW effect is advanced by the authors of this paper.

4.3 Lack of evidence of the AJW effect

Close examination of the empirical results provided in many empirical studies suggests that in many sectors the AJW effect could not be present since for the firms in these regulated industries, installed capital assets are complements to other inputs. In particular, network industries, which have very high capital to labour ratios under any regulatory regime due to the nature of the industry, are frequently characterised by complementarity of capital and labour. Recall that the AJW effect arises from the constraint the regulator places on the rate of return that accrues to installed capital. While it is possible that there may be some substitution possibilities between capital and labour prior to the acquisition of the capital, after installation the two inputs are likely to be complements. Sankar (1977) observes that

Dhrymes and Kurz [1964] found that, at the plant level, the partial elasticity of substitution between capital and labor is zero and that between capital and fuel is very small.... Further, using time series data for the U.S. electric utility industry for
the period 1949-68, Sankar [1972] found that an investment function based on a Leontief-type production model performed better than the functions based on a neoclassical model. Perhaps, a more realistic model would be the one which permits greater substitution possibilities before the investment is undertaken and less substitution possibilities after the investment is made. Sankar (1977) note 2, page 2.

And in Pescatrice & Trapani (1980) we find that

an analysis of the objectives of the private firms in the sample reveals some evidence consistent with internal cost minimization as predicted by the [AJW] regulatory model. However, not all of the perverse [distortionary] behavior predicted by the static rate-of-return model can be confirmed. In particular, the positive relation between nonbase input prices and their quantities demanded was not observed in most cases. The fact that these input demand distortions do not manifest themselves in the data is most likely due to the fact that the comparison was performed for the generation of power only and that substitution possibilities among inputs is somewhat limited by technology in this facet of the operation. Pescatrice & Trapani (1980), pages 274-275.

Finally, Nelson & Wohar (1987) in their study of US electric utilities note that

It is...possible to obtain three different estimates of the elasticities of scale and substitution for the unregulated technology. The traditional estimates are appropriate for the assumption that regulation is not binding, while the Fare-Logan estimates are consistent with the assumptions of binding regulation and cost minimization with respect to capital

[and, presumably, the assumption that the regulatory constraint is correctly specified].

The estimates...from the variable cost function are appropriate in both of these cases, and in the case when the firm is not in equilibrium with respect to it use of capital. Nelson & Wohar (1987), page 538.

It is worth observing that the estimate of the elasticity of substitution between capital and labour in this third specification – the variable cost function which is appropriate in all three cases – suggests that these two inputs are complements.²⁷

The three examples presented here are all drawn from electric utilities but most regulated network industries share fundamental technological characteristics with electricity – especially the high sunk cost of capital that is required for the network itself – and that’s why they are regulated,
usually. Thus the kind of industry which is most often regulated may also be the kind of industry in which we are least likely to see any significant AJW effect.

Consider telephone service, as another example. In commenting on the telecommunications industry Greenwald & Sharkey play down the need for concern over the AJW effect:

Another potential distortion due to rate of return regulation is the distortion in investment decisions which may lead to overly capital intensive technologies.... However, ...the size of the incentives involved is not clear, and the general empirical evidence supporting the existence of substantial economic inefficiencies of this sort is weak to non-existent. Greenwald & Sharkey (1989), page 325.

In a 1999 study of U.S. local telephone service Resende finds that, although ROR is supposed to cause input bias in the regulated firm according to the AJW model, the removal of ROR does not cause any productivity gain from an end to the hypothesized input bias:

The evidence indicates that alternative regulatory regimes (price-cap and incentive regulation) do not seem to play any role in improving productive efficiency, in comparison to traditional rate-of-return regulation. This result displays robustness with respect to the choice of the technical change variable. Resende (1999), page 41.

Conclusions

A few studies have found evidence of the AJW effect. Studies from the same period found no evidence and more recent papers have found no evidence of the AJW effect. Either there never was a very significant AJW effect and/or regulators read the economics literature, too, and took steps to mitigate the AJW effect.

Certainly, there is no justification for assuming the AJW effect in trying to reconstruct the unregulated cost or production function. Most regulated industries have been regulated for some time and technological change occurred in the regulated context. On the one hand, some regulatory constraints that appear to be binding are not; on the other, effects of regulation can persist even after
the regulatory constraint has been relaxed.\textsuperscript{28} If we assume that we can specify the regulatory constraint, assume it binds, and then proceed with estimation then we must accept that our estimation results are conditional on the accuracy of our specification of the regulatory constraint. There are two prescriptions that arise from this paper:

(1) For empirical research: Consider the AJW effect, if at all, as one only possible outcome among many in the context of a set of regulatory interactions. Do not assume that the AJW effect is binding without careful testing.

(2) For pedagogy: Reduce the emphasis of the AJW effect in the teaching of Industrial Organization and present it, if at all, as one only possible outcome among many in the context of a set of regulatory interactions, along with possible mitigating factors.

It is not that the search for the AJW effect has not produced some interesting papers, just that this search can be abandoned now for more productive enterprises. Explorers who set out to find the Northwest Passage made other interesting discoveries. “Research” into the nature of philogiston, the ether, and the philosopher’s stone yielded insights into chemistry, so too the exploration of the AJW effect has produced some important results for regulation economists (though neither immortality nor the secret of the transformation of lead into gold).

We can agree with Joskow (2005) that it is time to set aside further exploration which takes as its primary objective the identification of the size of AJW effect (this measurement could be “by-the-way” or “on-the-way-by”) but not that the research was “a waste of time and effort”.\textsuperscript{29} History has yet to decide.

\textsuperscript{28} See, e.g., Law & Nolan (2003).

\textsuperscript{29} It is not clear whether other researchers share my opinion. Joskow (2005) is cited once in the Social Sciences Citation Index but not in regard to the AJW effect. The citation occurs in Cetin & Oguz (2007).
References


Church, Jeffrey & Roger Ware (2000) Industrial Organization: A Strategic Approach (Toronto: The McGraw-Hill Companies, Inc.)


Figure 1
Profit Hill and Constraint Plane

Figure 2
(a) Impossible Location, Outside the Constraint Curve, of the Efficient Point $P$ for the Isoquant Through $K_{\text{max}}$
(b) Only Possible Location of $P$, Inside the Constraint Curve
Appendix 1, continued. Zajac (1970) Figures

![Diagram](image)

**Figure 3**
Possible Operating Points for the Firm
Appendix 2: Citations for Articles in Dataset

Papers published after 1989 ("in the last fifteen years" before Joskow’s 2005 paper) have a **bold** date of publication.

<table>
<thead>
<tr>
<th>No.</th>
<th>Citation</th>
</tr>
</thead>
</table>
Citations for Articles in “Dataset”, continued


Citations for Articles in “Dataset”, continued


Citations for Articles in “Dataset”, continued


Citations for Articles in “Dataset”, continued


Citations for Articles in “Dataset”, continued


Citations for Articles in “Dataset”, continued


Citations for Articles in “Dataset”, continued


