AD-E501 387

Copy 27 of 38 copies

AD-A236 571

IDA PAPER P-2407

A REGULATORY INTERPRETATION OF DoD PROFIT POLICY

Matthew S. Goldberg
Thomas P. Frazier
Thomas R. Gulledge, Jr.

October 1990

Prepared for
Office of the Assistant Secretary of Defense
(Program Analysis and Evaluation)

DTIC
ELECTE
JUN 0 3 1991
S B D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

IDA

INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

91-00895

91 5 30 062

IDA Log No. HQ 90-35590

DEFINITIONS

IDA publishes the following documents to report the results of its work.

Reports

Reports are the most authoritative and most carefully considered products IDA publishes. They normally embody results of major projects which (a) have a direct bearing on decisions affecting major programs, (b) address issues of significant concern to the Executive Branch, the Congress and/or the public, or (c) address issues that have significant economic implications. IDA Reports are reviewed by outside panels of experts to ensure their high quality and relevance to the problems studied, and they are released by the President of IDA.

Group Reports

Group Reports record the findings and results of IDA established working groups and panels composed of senior individuals addressing major issues which otherwise would be the subject of an IDA Report. IDA Group Reports are reviewed by the senior individuals responsible for the project and others as selected by IDA to ensure their high quality and relevance to the problems studied, and are released by the President of IDA.

Papers

Papers, also authoritative and carefully considered products of IDA, address studies that are narrower in scope than those covered in Reports. IDA Papers are reviewed to ensure that they meet the high standards expected of refereed papers in professional journals or formal Agency reports.

Documents

IDA Documents are used for the convenience of the sponsors or the analysts (a) to record substantive work done in quick reaction studies, (b) to record the proceedings of conferences and meetings, (c) to make available preliminary and tentative results of analyses, (d) to record data developed in the course of an investigation, or (e) to forward information that is essentially unanalyzed and unevaluated. The review of IDA Documents is suited to their content and intended use.

The work reported in this document was conducted under contract MDA 903 89 C 0003 for the Department of Defense. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

This Paper has been reviewed by IDA to assure that it meets high standards of thoroughness, objectivity, and appropriate analytical methodology and that the results, conclusions and recommendations are properly supported by the material presented.

Approved for public release, 7 March 1991; distribution unlimited.

UNCLASSIFIED

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
October 1990

3. REPORT TYPE AND DATES COVERED
Final Report, Oct 1989 – Oct 1990

4. TITLE AND SUBTITLE
A Regulatory Interpretation of DoD Profit Policy

5. FUNDING NUMBERS
C-MDA-903-89C-0003
T-Q7-665

6. AUTHOR(S)
Matthew S. Goldberg, Thomas P. Frazier, and Thomas R. Gulledge, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA-P-2407

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
OASD(PA&E)
Room 2D272, The Pentagon
Washington, D.C. 20301

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12A. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; 7 March 1991; distribution unlimited.

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Military procurement is bound by the Federal Acquisition Regulations (FAR). These regulations imply profit constraints that are similar, but not identical, to the constraints on electrical utilities studied in the economics literature. In addition, it is hypothesized that DoD contractors do not maximize short-run profit, but rather "hoard" labor by maximizing a utility function defined over profit and labor. Maximization of this utility function subject to the FAR implies certain restrictions on the elasticities of derived demand for input. These restrictions are tested using annual financial data on four large military aircraft manufacturers.

14. SUBJECT TERMS
Federal Acquisition Regulations (FAR), Profits, Department of Defense

15. NUMBER OF PAGES
38

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

IDA PAPER P-2407

A REGULATORY INTERPRETATION OF DoD PROFIT POLICY

Matthew S. Goldberg
Thomas P. Frazier
Thomas R. Gulledge, Jr.

October 1990

INSTITUTE FOR DEFENSE ANALYSES

Contract MDA 903 89 C 0003
Task T-Q7-665

PREFACE

This Paper was prepared by the Institute for Defense Analyses (IDA) for the Office of the Assistant Secretary of Defense (Program Analysis and Evaluation), under contract MDA 903 89 C 0003, Task Order T-Q7-665, issued 24 January 1989, and amendments. The objective of the task was to conduct an economic analysis of a selected set of contractors in the military aircraft manufacturing industry.

Howard J. Manetti served as the cognizant technical official for the task until his retirement in December 1989. Gary Bliss has been the technical official since then. This work was reviewed within IDA by Stanley A. Horowitz and David R. Graham. Thomas R. Gulledge, Jr., one of the authors of this paper, is an IDA consultant.

DTIC
COPY
INSPECTED

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

CONTENTS

TABLES

I. INTRODUCTION

The market in which the Department of Defense (DoD) procures military equipment is not fully competitive. Apart from foreign military sales, DoD is the sole purchaser of major items of military equipment. Moreover, the number of potential manufacturers of these items is often quite small as well.

The DoD applies a set of rules, known as the weighted guidelines, in determining the markups paid to manufacturers. The weighted guidelines are promulgated in the Federal Acquisition Regulations (FAR) [1]. In particular, the FAR allows for two components of markup above cost. One component is proportional to total allowable costs on the contract, and the other component is proportional to the net book value of the capital employed in production.

In light of the weighted guidelines and DoD's strong bargaining position, it appears fruitful to view production of military equipment as a regulated industry. The economics literature, beginning with Averch and Johnson [2], has devoted considerable attention to the behavior of regulated firms, particularly those in the electrical utilities industry. This literature has concentrated on the incentives that regulation provides the firms and on the efficiency of the firms' resulting behavior. In particular, it has been argued that the form of the regulation induces firms to "over-invest" in capital in an attempt to relax the regulatory profit ceiling.

A related strand of the economics literature has concentrated not on the regulatory constraints that apply to various industries, but rather on the behavioral objectives of the firms in these industries. Most prominently, Williamson [3] has developed a model in which firms "hoard" labor, hiring more workers than the number that would minimize the cost of producing the observed quantity of output.

Although hoarding of labor may appear inefficient from a short-run perspective, it may indeed be efficient from a long-run perspective. It can be quite expensive to lay off workers in response to a temporary decline in business, only to rehire those workers when business returns to its normal, long-run level. A more efficient strategy may be to retain workers even in periods when they cannot be fully utilized.

These considerations are particularly compelling for firms in the defense industry. Temporary declines in business are quite common; these declines may prevail industry-wide, or may apply to individual firms that lose specific contract competitions. Moreover, it is extremely expensive to lay off and subsequently rehire skilled workers in this industry, especially engineers.

In this paper, we draw inspiration from the economics literature, and develop a mathematical model that purports to describe the behavior of firms in the defense industry. Under this model, firms again hoard labor by hiring more workers than the number that would be efficient in the short-run. In addition, firms are constrained by an equation that represents the effect of the FAR regulations on their profit margins.

The conjunction of labor hoarding and the profit constraint leads to certain restrictions on the firm's demand curve for labor. These restrictions differ dramatically from those implied by the more conventional model in which firms simply minimize cost. In principle, a detailed examination of the demand curve for labor permits discrimination between these two models.

In Section II of this paper, we review some models from the economics literature, and adapt them for application to the defense industry. Section III contains a review of the empirical tests that have been applied to this class of models in the economics literature. In Section IV, we develop a stricter testing procedure that will be applied to discriminate between our new model and the conventional model of cost minimization. A description of the data that will be used to conduct the empirical tests is given in Section V.

In Section VI, we report the outcome of the empirical tests, using data from four large aerospace manufacturers. It will be seen that the restrictions implied by our model are generally supported by the data. Finally, Section VII contains the conclusions of the analysis.

II. BEHAVIORAL MODELS

A. NOTATION

The firm produces a single output, denoted Q, using inputs of labor, L, capital, K, and materials, M. The production function is denoted Q(L, K, M). Revenue is given by the function $R(Q) = R[Q(L, K, M)]$. The input prices are P_l, the wage rate of labor; P_k, the ownership cost of capital; and P_m, the price of materials. Profit is defined as revenue minus cost, $\pi = R(Q) - C = R(Q) - P_lL - P_kK - P_mM$.

B. WILLIAMSON MODEL

Williamson [3] proposed a model in which the firm's managers derive utility not only from profits, but also from the firm's expenditures on corporate staff. This model suggests a utility function of the form $U(\pi, P_lL)$. Edwards [4] and Hannan [5] have conducted empirical tests of Williamson's model, using data on commercial banks. They both concluded that managers' preferences for labor are manifested by hiring more workers, rather than by paying inflated wages to a fixed number of workers.[1] Therefore, the second argument of the utility function may be replaced by the amount of labor employed, simplifying the utility function to $U(\pi, L)$.

Further, it is easy to show that maximizing $U(\pi, L)$ in turn implies maximizing L subject to the constraint $\pi \geq \pi^*$. We may interpret the value π^* as the minimum amount of profit that the firm's stockholders demand; lower profit than π^* would lead the stockholders to expel the current management.

C. AVERCH-JOHNSON MODEL

In describing the electrical utilities industry, Averch and Johnson [2] proposed a model in which the firm maximizes profit, subject to the constraint that the "return on

[1] For example, Hannan ([5], p. 894) states: "The identical impact of concentration on [labor expenses] and [the number of workers] is consistent with Edwards' finding that management indulges its taste for expenses primarily by hiring excess staff rather than by paying higher salaries."

capital" be at most a specified percentage of the capital stock. This percentage, denoted s, is assumed greater than the cost of capital, P_k. The constraint may be written as:

$$(1) \qquad R[Q(L,K,M)] - P_l L - P_m M \leq sK \ ,$$

or equivalently:

$$(2) \qquad \pi \leq (s - P_k)K \ .$$

Note that the restriction $s \geq P_k$ allows the regulated firm some positive profits. The allowed amount of profit is an increasing function of K, leading the firm to invest in additional capital in an attempt to raise the profit ceiling. This is the famous "over-capitalization" result of Averch and Johnson.

D. ADAPTATION OF REGULATORY CONSTRAINT FOR DOD

The "return on capital" on the left-hand side of Equation (1) does not include any interest costs. The exclusion of interest costs may indeed be appropriate for the application of this model to the electrical utilities industry. However, the issue at hand is whether the exclusion is equally appropriate for DoD contractors.

At first blush, the exclusion of interest costs appears to be consistent with the long-standing DoD policy prohibiting reimbursement of interest expenses. According to Osband ([6], p. 15):

> Ever since the first set of formal cost principles was issued in 1940, the Government has explicitly disallowed interest charges. That is, not only is no markup calculated on interest costs, but the very interest itself is not reimbursed. It accrues as a wasteful expense, to be subtracted from the nominal calculated profit. Government justifications for not allowing interest include discouragement of excessive debt financing, avoidance of disputes over appropriate financing costs, and neutralization of special competitive advantages of cash-rich big businesses [sic].

Although DoD does not allow interest as a reimbursable expense, it has since 1977 allowed interest charges in the computation of "profit." While the DoD accounting definitions of "cost" and "profit" differ from those advanced by economists, the end result is that DoD contractors are compensated quite generously for the costs of capital ownership.

Specifically, among the components of profit that DoD pays its contractors are the *facilities capital cost of money* and the *facilities capital markup*, both introduced in 1977. In each case, the net book value of capital employed in production is multiplied by a markup rate, and the result is summed for each year of project duration. This procedure is applied without regard to whether the contractor's source of funds is equity or borrowed capital.

The markup rate used to compute the facilities capital cost of money is known as the treasury rate. Rogerson [7] shows that the treasury rate is generally one percentage point higher than the imputed interest rate on U.S. government bonds with a maturity of five years. The extra percentage point is presumably a risk premium, reflecting the fact that corporations borrow at a higher interest rate than does the government.

The facilities capital markup is an additional component of profit, presumably compensating for the loss of liquidity when corporations invest in physical rather than financial assets. The current markup rates are given by the ranges 10 to 20 percent per year for buildings, and 20 to 50 percent per year for equipment. The exact values applied to any particular contract are the result of negotiation between the contractor and the DoD contracting officer.

We view the facilities capital cost of money as compensation for the opportunity cost of capital, $P_k K$, and the facilities capital markup as pure profit, $s_1 K$. Hence the contractor's revenues from DoD consist of these two quantities, plus direct reimbursement for expenditures on labor and materials.

$$R[Q(L, K, M)] \leq P_l L + P_k K + P_m M + s_1 K \ , \tag{3}$$

or equivalently:

$$\pi \leq s_1 K \ . \tag{4}$$

From Equation (4), it appears that the more appropriate version of the regulatory constraint allows deduction of interest costs in computing the "return on capital."

Finally, Equation (4) must be augmented to include the remaining components of profit that DoD pays its contractors. These components are each computed as a percentage of total allowable costs, as indicated in Table 1. The first column of the table simply names the various components of profit. The DoD contracting officer selects a profit rate for each component, which must lie between the lower and upper limits indicated in Table 1. The table also indicates the so-called "normal" profit rate, which is just the midpoint of the lower and upper limits. If, for example, the contracting officer determines that the project contains an unusual amount of technical risk, then he is empowered to offer the upper limit of a 1.8-percent markup on this component. The profit rate selected is then applied to the base indicated in the final column of the table. The cost-based components are proportional to total costs minus General and Administrative (G&A) costs.

Table 1. DoD Profit Policy as of 1987

	Allowable Range			
Component of Profit	Low	Normal	High	Base to which Applied
Technical Risk	0.6%	1.2%	1.8%	Total Cost – G&A
Management Complexity	0.6%	1.2%	1.8%	Total Cost – G&A
Cost Control	0.8%	1.6%	2.4%	Total Cost – G&A
Contract Risk				
Firm fixed price	2.0%	3.0%	4.0%	Total Cost – G&A
Fixed price incentive	0.0%	1.0%	2.0%	Total Cost – G&A

In practice, the DoD contracting officer and the contractor negotiate over the total profit rate, not the individual components of profit. This practice is condoned by the FAR regulations[2]: "Specific agreement on the exact values or weights assigned to individual profit-analysis factors is not required during negotiations and *should not be attempted.*" [Emphasis added.]

There has historically been little variation in the profit rates assigned to the cost-based components. A study by the Logistics Management Institute [8] analyzed profit margins on 3,686 manufacturing contracts negotiated over the period 1980-1982. The markup rate on cost (i.e., the sum of the cost-based components of profit, divided by contract cost) had a sample mean of 11.5 percent and a standard deviation of only 2.9 percent. Hence there was little variation in the markup rate on cost, either across contracts or across the three years studied.

Let s_2 denote the markup rate on cost (e.g., $s_2 = .115$). Then the final form of the regulatory constraint is:

$$\pi \leq s_1 K + s_2 C \quad . \tag{5}$$

2 See Reference [1], section 15.807.

E. HYBRID MODEL

We now combine various aspects of the Williamson and Averch-Johnson models, to arrive at a hybrid model that we propose to describe the behavior of DoD contractors. First, like Williamson, we assume that the firm maximizes labor subject to the constraint $\pi \geq \pi^*$. One interpretation of this formulation is that firms "hoard" labor, retaining workers beyond the point that maximizes profit. Another interpretation is that firms are in fact maximizing long-run profit. To do so, firms refrain from laying off workers in response to short-run fluctuations in product demand. The hoarding of labor, although apparently sub-optimal in the short-run, may actually be consistent with profit maximization in the long-run.[3]

This hypothesis may be sharpened, based upon conversations with industry experts. These experts contend that, because hiring costs are much steeper for engineers than for production labor, only the engineering component of labor is hoarded. We therefore partition the total workforce into engineering labor (L_e) and production labor (L_p), with respective prices P_e and P_p. Our revised hypothesis is that the firm maximizes L_e subject to the constraint $\pi \geq \pi^*$.

Recall that the constraint $\pi \geq \pi^*$ is imposed by the firm's stockholders, who demand at least a minimum amount of profit. However, profits are constrained in the opposite direction by DoD, in accordance with Equation (5). Therefore, our final hypothesis is maximization of L_e, subject to the two contraints:

(6) $$\pi \geq \pi^*, \qquad \pi \leq s_1 K + s_2 C \ .$$

F. PROPERTIES OF HYBRID MODEL

Our hybrid model implies several restrictions on the firm's demand function for engineering labor. Recall that total cost is equal to: $C = P_eL_e + P_pL_p + P_kK + P_mM$. Using this definition and assuming that the two constraints in Equation (6) hold as equalities, we find:

(7) $$\pi^* = s_1 K + s_2 C = s_1 K + s_2 (P_eL_e + P_pL_p + P_kK + P_mM) \ .$$

Solving this equation for L_e gives the firm's demand function for engineering labor:

(8) $$L_e = [\pi^* - (s_1 + s_2 P_k) K - s_2 (P_pL_p + P_mM)]/(s_2 P_e) \ .$$

---

[3] The suggestion that labor hoarding may be efficient in the long-run was made by Miller [9].

The partial derivative of L_e with respect to P_e is equal to:

(9) $$dL_e/dP_e = [(s_1 + s_2 P_k) K + s_2 (P_p L_p + P_m M) - \pi^*]/(s_2 P_e^2) \ .$$

The elasticity of L_e with respect to P_e is defined as $e(L_e, P_e) = (P_e/L_e)\, dL_e/dP_e$, and the share of engineering labor in total cost is defined as $C_e = P_e L_e/C$. Given these definitions, we find:

(10) $$e(L_e, P_e)/C_e = [(s_1 + s_2 P_k) K + s_2 (P_p L_p + P_m M) - \pi^*]/[P_e^2 L_e^2 s_2/C] \ .$$

Although this expression seems formidable, it follows from equation (8) that the numerator is simply $-s_2 P_e L_e$. Hence Equation (10) reduces to:

(11) $$e(L_e, P_e)/C_e = -C/(P_e L_e) \ .$$

We may also derive the elasticities of demand for engineering labor with respect to the prices of production labor, capital, and materials. For example, the partial derivative of L_e with respect to P_k is $dL_e/dP_k = -K/P_e$, yielding the result:

(12) $$e(L_e, P_k)/C_k = -C/(P_e L_e) \ .$$

Similar results are obtained for the prices of materials and production labor, yielding the overall result:

(13) $$e(L_e, P_e)/C_e = e(L_e, P_p)/C_p = e(L_e, P_k)/C_k = e(L_e, P_m)/C_m \ .$$

Economists classify a pair of inputs as "substitutes" if an increase in the price of the first leads to an increase in purchases of the second, and "complements" in the event of a decrease in purchases of the second. The "own-price" effect, $e(L_e, P_e)$, is always negative. Equation (13) implies that the "cross-price" effects are negative as well, so that production labor, capital, and materials are *complements* with engineering labor, not substitutes.

The intuition behind Equation (13) is as follows. If the stockholder constraint is binding, then the firm is earning profit of exactly π^*. If the price of *any* input increases and the firm does not adjust, its profit will fall below π^*, violating the stockholder constraint. In order to restore profit of π^*, the firm must curtail its hoarding of labor. Hence L_e will decrease in response to an increase in *any* input price, and the cross-price effects on demand for engineering labor must all be negative.

Our model stands in sharp contrast to the conventional model of cost minimization. In that model, it can be shown that the elasticities of demand satisfy the following restriction[4]:

(14) $$e(L_e, P_e) + e(L_e, P_p) + e(L_e, P_k) + e(L_e, P_m) = 0 \ .$$

Because $e(L_e, P_e)$ is always negative, at least one of the cross-price effects must be positive to maintain Equation (14). Therefore, the conventional model implies that either production labor, capital, or materials must be a substitute for engineering labor; our model implies instead that these inputs are all complements for engineering labor. This discrepancy between the predictions of the two models may be used to distinguish them empirically.[5]

---

4 See, for example, Chambers ([10], p. 65).

5 It may be shown, by more advanced methods, that Equation (13) holds even when engineering labor is maximized subject to the stockholder constraint alone (i.e., even when the regulatory constraint is not binding). These restrictions, although apparently not recognized in the literature, could have been used to test the basic Williamson model.

III. HISTORY OF EMPIRICAL TESTING

A. AVERCH-JOHNSON MODEL

Averch and Johnson published their theory of the regulated firm in 1962. Serious empirical tests of their theory were not published until twelve years later. All of these tests used data from the electrical utilities industry. The seminal papers were by Robert Spann [11] and Leon Courville [12], published in the Spring 1974 issue of the *Bell Journal of Economics and Management Science*.

Spann estimated a production function for electricity so that the variables on the right-hand side were input quantities. Unfortunately, input quantities are endogenous variables, chosen by the firm in an effort to minimize cost or maximize profit. It is well-known that least-squares estimates of regressions on endogenous variables are biased.

Spann's procedure purported to estimate the Lagrange multiplier associated with the rate-of-return constraint. This multiplier was treated as a single, fixed number. However, the multiplier is more properly considered as varying across both firms and time periods, depending on the "tightness" of the regulatory constraint at each instant. Therefore, the single estimate of the multiplier is at best an average of many underlying values. Notwithstanding these criticisms, Spann's estimate of the multiplier was significantly positive, lending support to the Averch-Johnson hypothesis.

Courville also estimated production functions for electricity, hence his regressions were biased as well due to endogenous variables on the right-hand side. Again, criticisms notwithstanding, Courville found statistical evidence in support of the Averch-Johnson hypothesis.

Petersen [13] presented yet another empirical test, one year later in the same journal. He estimated a cost function rather than a production function. The variables on the right-hand side were input prices (not quantities) and the quantity of output (measured in kilowatt hours of electricity). Input prices are exogenous if the firm is one of many purchasers of each input, because in that situation the firm is too small to influence the price. However, the quantity of output is an endogenous variable, again chosen by the firm in an effort to maximize profit. Notwithstanding this criticism, Petersen found statistical evidence in support of the Averch-Johnson hypothesis.

Boyes's [14] estimated input demand functions, *conditional* on the quantity of output. Like Petersen, the variables on the right-hand side were input prices and the quantity of output, the latter an endogenous variable. Like Spann, he estimated a single value of the Lagrange multiplier associated with the rate of return constraint, ignoring likely variation in the multiplier across firms and time periods. He could not reject the hypothesis that the Lagrange multiplier equals zero, so the rate of return constraint is not binding on the firm. Boyes's the only major study to find evidence against the Averch-Johnson hypothesis in the electrical utilities industry.

The definitive study was written by Cowing [15]. He estimated the *unconditional* input demand and profit functions. In deriving these functions, the output quantity is chosen along with the input quantities, so the only remaining arguments of the functions are the input and output prices and the mandated rate of return. All of these arguments are considered exogenous, hence there is no bias due to endogenous variables.

Cowing performed three distinct tests of the Averch-Johnson hypothesis. First, several terms in the input demand and profit functions involve the mandated rate-of-return. Cowing tested whether these terms were all simultaneously zero, which would imply that the regulation was ineffective. He rejected this hypothesis for all three time periods examined, hence concluding that the regulation was indeed effective.

Second, Cowing is the only author to test the well-known (even at that time) restriction of the Averch-Johnson model, $dK/dP_k=0$. He notes on p. 231: "An additional test of general regulatory effectiveness . . . follows from noting that the [Averch-Johnson] model of the regulated firm implies . . . $dK/dP_k=0$." Using a likelihood ratio test, he accepts this restriction in two of the three time periods. Using the Wald test (i.e., comparing the estimate of dK/dP_k to its standard error), he accepts this restriction in only one time period, and even then by a small margin (the t-statistics are 1.78, 1.99, and 3.26). As he notes, the two test criteria are asymptotically equivalent, but may yield divergent results in finite samples.

Finally, Cowing is the only author to estimate a separate value of the Lagrange multiplier for each firm and time period. The multiplier was significantly positive for 1 of 21 firms during the period 1947-1950, 12 of 26 firms during the period 1955-1959, and 17 of 23 firms during the period 1960-1965 (the period 1951-1954 was deleted because the estimation algorithm did not converge).

In summary, the evidence indicates that rate-of-return regulation was effective in the electrical utilities industry, at least for some firms and some time periods.

B. WILLIAMSON MODEL

Williamson published his theory of the firm in 1964. However, it was Edwards [4] who performed the first serious empirical test of Williamson's theory. Edwards' idea was quite simple and ingenious. A profit-maximizing monopolist restricts output, charging a higher price and selling fewer units of output than would a competitive industry facing the identical demand curve. Hence a monopolist will normally purchase fewer units of each input as well.[6] Consider, however, a monopolist whose utility function includes the quantity of labor employed in addition to the level of profit. This monopolist may actually employ more labor than would a competitive industry, if the effect of labor on utility outweighs the tendency to restrict output. Note the implicit assumption that only the monopolist, being insulated from competitive pressures, has the latitude to indulge his preference for labor.

Using data on the banking industry, Edwards estimated the demand function for labor. Among the variables on the right-hand side, he included a dummy variable equal to 1.0 if the concentration ratio exceeds a threshold value.[7] If the coefficient of the dummy variable is positive, then the effect of labor on utility is definitely present, because it outweighs the offsetting tendency for monopolists to restrict output.

In addition to the dummy variable for concentration, a demand function for labor should include the prices of labor and all other inputs. Edwards did include the price of labor, but not the price of capital (i.e., the interest rate at which banks themselves may borrow money). He justifies this omission on p. 155, "I assumed that the cost of capital was identical for all banks, a common assumption in banking studies." Clearly, one cannot include in a regression a variable that takes the same value at all data points.

Edwards' claim of common interest rates may be valid, particularly if all banks borrow money in the same national market. Unfortunately, however, this approach precludes testing the proportionality restrictions on labor demand that are derived in this paper (i.e., the proportionality between the elasticities of labor demand with respect to the

---

6 We say "normally" to admit the possibility of so-called inferior inputs, which are used in smaller quantities as output expands and larger quantities as output contracts. A contemporary example of an inferior input might be a 286-computer, which is replaced by a 386-computer when output expands. The possibility of inferior inputs, first raised by Bear [16], was implicity ruled out by Edwards's choice of functional form.

7 The concentration ratio is defined as bank deposits of the three largest banks, divided by total bank deposits in the Standard Metropolitian Statistical Area (SMSA). A larger concentraction ratio indicates greater monopoly power on the part of the largest banks.

price of labor and the price of capital). Neither Edwards nor any of his followers seemed aware of these restrictions, although they probably could not have tested them in any case using data on the banking industry.

Edwards found strong evidence that monopolistic banks employ more labor, thus implying that their utility functions include the quantity of labor as well as the level of profit. In addition, he found that the preference for labor is manifested by hiring more workers, rather than by paying inflated wages to a fixed number of workers.

Edwards's findings were replicated in a number of subsequent studies. Hannan [5] used data on individual banks rather than aggregate data on SMSAs. He too omitted the price of capital, stating on p. 894, "Under the assumption that the cost of capital is the same for all banks, a proxy for [the cost of capital] is not included in the estimation." His results confirm those of Edwards; in particular, he finds an effect of concentration on the number of workers but not on the average wage per worker.

Rhoades [17] examined various expense categories, apart from labor expense, that monopolists might expand in an effort to maximize utility. These expense categories included such sundry items as furniture, office supplies and stationery, charitable donations, books and periodicals, dues and memberships, and travel and entertainment. Among this laundry list of expense categories (inexplicably, laundry expense was not analyzed), only charitable donations and dues and memberships were positively related to the concentration ratio.

Hannan and Mavinga [18] questioned the assumption that monopoly power is the correct measure of management's latitude to indulge its preference for labor. Instead, they argue that managers have more freedom when the ownership of the firm is dispersed among a large number of stockholders, with no single block of stockholders owning a significant share of the total stock. They label this situation as "management-control" of the firm in contrast to "owner-control." Their findings, again for the banking industry, are that total wage and salary expenses are higher in situations where the firm is both management-controlled and possesses some monopoly power.

Smirlock and Marshall [19] argue that monopoly power must be present not in conjunction with management control, but rather in conjunction with firm size. Presumably, stockholders find it more difficult to monitor management behavior in a large firm, offering managers greater freedom to indulge their preferences. They find that firm size (measured by bank assets) is a better predictor of labor per unit output than is the concentration ratio. Indeed, once firm size is taken into account, the concentration ratio no longer has a significant effect on the quantity of labor employed. These findings do not contradict Williamson's hypothesis, but they do suggest that firm size rather than monopoly power is the factor that enables management to indulge its preferences.

Finally, Awh and Primeaux [20] tested Williamson's hypothesis using data on the electrical utilities industry. However, they were unable to obtain data on either total wages and salaries or total employment for most of the firms in their sample. Instead, they used "sales and administrative expenses" as the object of managerial interest. Their failure to obtain data on total employment is somewhat disappointing, because both Boyes [14] and Cowing [15] were able to obtain these data in their respective studies of the electrical utilities industry. Awh and Primeaux found that sales and administrative expenses are, if anything, lower in monopolistic situations than in competitive situations. Although they interpret this finding as evidence against Williamson's hypothesis, this conclusion is predicated upon their dubious choice of sales and administrative expense as the dependent variable.

C. HYBRID MODELS

Two studies have attempted to combine aspects of the Averch-Johnson and Williamson models. Neither study reports any empirical results, but both studies offer interesting theoretical analysis.[8] Crew and Kleindorfer [23] formulated a model in which the firm maximizes utility, a function of profit and "staff expenditure," subject to a rate-of-return constraint. Unfortunately, they obtained very few analytic results, relying instead upon a numerical simulation to determine the properties of their model.

A similar model was analyzed by Arzac and Edwards [24]. They argue that the over-capitalization result of Averch-Johnson may be offset by the tendency for monopolistic managers to hire excessive amounts of labor, possibly leading to an efficient capital/labor ratio. They state on p. 48:

> If [preference for labor expense is operative] and regulation is effective, the tendency of the unregulated expense-preference firm to use too little capital must be balanced against the tendency for rate-of-return regulation to cause firms to use too much capital. Indeed, these two forces may totally offset one another, so that there is no internal inefficiency at all. . . . Thus, regulation may actually make firms more efficient internally.

---

[8] Two other studies, Bailey and Malone [21] and McNicol [22], have combined the Averch-Johnson model with a behavioral objective of revenue-maximization. Again, no empirical results were reported.

Although the Arzac-Edwards analysis is interesting, they fail to notice the proportionality restrictions on labor demand that are derived in the current paper. However, they do seem aware that the proper approach for sorting out the various models involves careful estimation of the input demand functions. They state on p. 50:

> In principle, therefore, a complete test of factor usage in regulated firms could distinguish whether such firms were profit maximizers. . . . To date, none of the [Averch-Johnson] studies have estimated a complete enough system of factor demand functions to enable us to distinguish the [Averch-Johnson] hypothesis from the managerial discretion hypothesis.

IV. METHODOLOGY

A. MODEL SPECIFICATION

In order to test the proportionality restrictions in Equation (13), we must estimate the elasticities of demand for engineering labor with respect to the price of engineering labor and the prices of all other factor inputs. Hence the quantity of engineering labor is the appropriate left-hand variable in a regression model. Among the right-hand variables, we must include the prices of engineering labor, P_e; production labor, P_p; capital, P_k; and materials, P_m.

We specified a log-linear relationship between the quantity of engineering labor and the various factor prices. We also included output quantity, a technology index, plus three dummy variables to allow separate intercepts for each of the four firms in the sample. The regression model is:

$$(15) \qquad \ln(L_e) = b_0 + b_1 \ln(P_e) + b_2 \ln(P_p) + b_3 \ln(P_k) + b_4 \ln(P_m) + b_5 \ln(Q) + b_6\,\mathrm{TECH} + b_7 D_1 + b_8 D_2 + b_9 D_3 + e_t \ .$$

We assumed a first-order autocorrelation structure for e_t: $e_t = \rho e_{t-1} + v_t$, where v_t is distributed independent normal.

The log-linear specification was chosen for the following reasons. First, under this specification the coefficients on the factor prices may be interpreted as elasticities. This feature is attractive because the proportionality restrictions apply directly to the elasticities. Second, as we will see, the R-squared statistic for the estimated model is extremely high. The excellent fit obviates the need to introduce higher-order terms (i.e., squares and cross-products of the logarithmic prices).

B. HYPOTHESIS TESTING

This section will outline the statistical procedure for testing the proportionality restrictions. Equation (13) constitutes three independent restrictions, which may be expressed in linear form upon cross-multiplication:

$$(16) \qquad T_1 = C_p e(L_e, P_e) - C_e e(L_e, P_p) = 0 \ ,$$

$$(17) \qquad T_2 = C_k e(L_e, P_e) - C_e e(L_e, P_k) = 0 \ ,$$

(18) $T_3 = C_m e(L_e, P_e) - C_e e(L_e, P_m) = 0$.

As noted earlier, the elasticities in Equations (16) through (18) correspond to the regression coefficients in a log-linear specification. These regression coefficients are estimated from finite samples, and are subject to sampling variation. The magnitude of the sampling variation is expressed by the standard errors, which are available from the regression output.

We must also supply the cost shares in Equations (16) through (18). The cost shares vary by firm, so that the restrictions may be valid for one firm but invalid for another. Therefore, we endeavored to test the restrictions separately for each firm. We estimated the cost shares for each firm as the simple averages over the years observed in the sample.

Strictly speaking, sampling variation in the average shares introduces additional randomness into T_1 through T_3, beyond that embodied in the regression coefficients. We choose to ignore this variation, instead treating the shares for each firm as fixed constants. In so doing, we underestimate the total variance in T_1 through T_3, making it more likely that we will reject the proportionality restrictions. In effect, we bias the test against our model; if the proportionality restrictions hold despite this bias, then we can have additional confidence in their validity.

A conventional procedure in hypothesis testing is to estimate the model twice, once with the restrictions imposed and once without the restrictions. A comparison between the value of some criterion function (e.g., sum of squared deviations) in the two situations provides the basis for either accepting or rejecting the hypothesis.

Unfortunately, this procedure could not be applied using our data. If there were enough years of observation for each firm, a separate regression model could be estimated for each firm. The model for each firm could be estimated both with and without the restrictions, using the firm-specific cost shares in the former case. A comparison of the sum of squared deviations would determine acceptance or rejection, separately for each firm.

The difficulty arises because we do not have enough years of observation to estimate a separate regression model for each firm. Instead, we estimate a combined regression model, with separate intercepts (captured by the dummy variables) but common slopes. Because the cost shares are firm-specific, no single restriction of the form represented by Equations (16), (17), or (18) could properly be applied to the entire sample.

Even if the model were valid, a restriction that applies to one firm would not apply to the other firms in the sample if the cost shares were different.

To avoid this difficulty, we estimated the combined regression model only once, without imposing any restrictions. We then evaluated the linear combinations T_1, T_2, and T_3 separately for each firm, using the common elasticities but the firm-specific cost shares. In effect, we obtained linear combinations, T_{1j}, T_{2j}, and T_{3j} for firms j = 1, 2, 3, 4. For firm j, we then tested the hypothesis that T_{1j}, T_{2j}, and T_{3j} are *simultaneously* equal to zero. The results of this test could conceivably vary by firm, if the cost shares lined up in proportion to the elasticities for one firm but not for another.

Finally, we require a criterion for determining whether the sample values of T_{1j}, T_{2j}, and T_{3j} for firm j are significantly different from the hypothesized value of zero. Because we are testing a joint hypothesis on three linear combinations, the conventional significance levels do not apply. For example, if T_{1j} were equal to 1.96 (or -1.96), we would reject the hypothesis $T_{1j} = 0$ at the .05 significance level. Similarly, if T_{2j} and T_{3j} were equal to 1.96 (or -1.96), we would reject the hypotheses $T_{2j} = 0$ and $T_{3j} = 0$ at the .05 significance level. However, the overall significance level of the *joint* hypotheses is not equal to .05 (nor is it equal to .15).

Recall the definition of significance level: the probability that, because of sampling variation, the hypothesis is rejected, *despite the hypothesis being true*. Suppose we reject the joint hypothesis $T_{1j} = T_{2j} = T_{3j} = 0$ whenever the sample value of either T_{1j}, T_{2j}, or T_{3j} exceeds 1.96 in absolute value. If the joint hypothesis is true, the probability of each of these events, taken individually, is .05. The overall significance level is the probability that *either* T_{1j}, T_{2j}, or T_{3j} exceeds 1.96 in absolute value.

Computing the overall significance level in this situation is known as the problem of *multiple comparisons*. We demonstrate in the appendix that the overall significance level is at least equal to the maximum significance level of the three tests, and at most equal to the sum of the three significance levels. In the example above, the overall significance level is bounded between .05 and .15.

If the three tests were statistically independent, then the overall significance level would equal the complement of the probability that none of the tests leads to rejection: $1 - .95^3 = .1426$. However, the three tests for each firm are conducted using overlapping sets of elasticities, computed from a common data sample. Because independence clearly fails in this case, we will express all significance levels in terms of lower and upper bounds. These bounds are also known as the *Bonferroni inequality*. Fortuitously, the Bonferroni bounds will turn out to be quite tight in our application.

V. DATA

The data were provided by four large aircraft manufacturers. Information was collected not at the corporate level, but specifically at the level of the plants or divisions that produce military aircraft. There are a total of 66 annual data points. The data series ends in 1987 for all four firms, and begins in 1970 for two of the firms, 1972 for the third, and 1974 for the fourth.

The data have been adjusted and normalized to account for changes in organization and accounting systems. All variables are measured in thousands of 1987 dollars, using deflators that will be described in this section. The variables may be grouped into four categories: input quantities, input prices, output quantity, and a measure of product technology.

A. INPUT QUANTITIES

The workforce in each firm was partitioned into three categories. The first category consists of workers in the "occupancy pool" who perform maintenance on facilities and equipment. This category of workers is best viewed as contributing to the services of capital, rather than as labor. The cost of these workers will be included as a component in the price of capital, discussed below.

All remaining workers were partitioned into either engineering labor (L_e) or production labor (L_p). These workers were classified on the basis of work center (i.e., design versus production) rather than occupation. It is conceivable that some craftsmen were misclassified as engineering labor (e.g., those who worked on scale models in the design center), and some engineers were misclassified as production labor (e.g., operations researchers who worked on scheduling problems in the production center). For the most part, however, we expect the classification by work center to agree with the worker's occupation. Finally, all quantities were expressed in full-time equivalent man-years.

B. INPUT PRICES

Our analysis requires data on the prices of the productive factors: capital, engineering labor, production labor, and materials. We view the price of capital as the annual dollar cost per dollar of capital stock:

$$(19) \qquad P_k = [(\text{Depreciation} + \text{Utilities} + \text{Taxes} + \text{Maintenance})/\text{Net Book Value}] + \text{Normal rate of return}$$

Except for the normal rate of return, all of the components in Equation (19) were supplied by the contractors. In particular, the maintenance component represents the labor costs of the workers who performed maintenance on facilities and equipment. We approximated the normal rate of return using Moody's Aaa corporate bond rate, deflated by the GNP implicit price deflator [25].

The prices of engineering labor (P_e) and production labor (P_p) were also supplied by the contractors. These prices are measured as the CPI-adjusted, average annual cost of wages plus fringe benefits. Finally for the price of materials (P_m) we used the index for aircraft materials in Standard Industrial Classification (SIC) 3721. The use of an index to measure price trends follows the precedent set by Evans and Heckman [26].

C. OUTPUT QUANTITY

Our data set does not contain a measure of physical output rate, aggregated across aircraft types within each plant and year. Instead, we constructed a measure of value-added, defined as total cost minus the sum of direct materials, subcontracting, and General and Administrative costs.

We were initially concerned that value added might be "endogenous," chosen by the firm in an effort to maximize profit, utility, or some similar objective. An endogenous variable induces reverse causation and biased coefficient estimates. However, we will show later using Hausman's test [27] that endogeneity was not in fact a problem. This test requires an instrumental variable for value added. To construct the instrumental variable, we note that current activity in a plant is related to aircraft that will be delivered in the current year or in the next few years. We thus regressed value added on current deliveries and deliveries in the next two years. The two-year horizon was selected because it is consistent with known aircraft production profiles. Finally, the prediction equation for value added contained a first-order correction for autocorrelation.

D. TECHNOLOGY MEASURE

Product technology in the aerospace industry has been changing over time. To control for the effects of changing technology, a technology variable was constructed for inclusion in the regression equations. Company delivery schedules were examined, and data were collected on the types of aircraft under construction in each plant in each year. For each type of aircraft, the following index was computed[9]:

$$T_i = \left(\frac{1}{A}\right)\left(\frac{\text{EMW-ENW}}{\text{STW}}\right) \quad , i = 1, \ldots, J \tag{20}$$

A = percent aircraft aluminium content

STW = aircraft structure weight

EMW = aircraft empty weight

ENW = aircraft engine weight

J = the number of aircraft types in the contractor's plant in a given year.

The technology index for each plant in each year is a linear combination of the relevant values computed in Equation (20). The weights for the linear combinations (W_i) are proportional to the total number of each type of aircraft in the contractor's plant in each year. Therefore, the index for a given contractor in a given year is:

$$\text{TECH} = \sum_{i=1}^{J} W_i T_i \quad . \tag{21}$$

This index attributes higher technology to aircraft with a lower aluminum content, and a correspondingly higher content of advanced materials. The index also attributes higher technology to aircraft with greater "density," i.e., a higher percentage of non-engine (e.g., avionics) weight. Our index is preferable to using a uniform time trend for all firms, because the latter would ignore aircraft type.

---

9 The index was suggested by Bruce Harmon, and the data for its construction were taken from his study of aircraft development costs [28].

VI. EMPIRICAL FINDINGS

We first perform the Hausman [27] test to determine whether value added is endogenous. To perform this test, we augmented equation (15) to include the instrumental variable for value added as well as value added itself (both expressed in logarithms). Although we do not report the entire regression, we note that the instrumental variable is insignificant with a t-statistic of only 0.108. We infer from this result that value added is not endogenous.

The estimate of Equation (15), using value added, is found in Table 2. The model fits quite well, with an R-squared statistic of .985. The own-price effect is negative and statistically significant. Production labor and materials are indeed complements to engineering labor, although neither of these effects is statistically significant. However, capital appears to be a substitute for engineering labor, and this effect is statistically significant.

Table 2. Regression Estimates: Logarithmic Demand for Engineering Labor

Variable	Coefficient	Standard Error	T-Statistic
Intercept	-0.206	0.854	-0.241
ln (Engineering Labor Price)	-0.531*	0.253	-2.094
ln (Production Labor Price)	-0.231	0.301	-0.769
ln (Capital Price)	0.136*	0.053	2.587
ln (Materials Price)	-0.074	0.082	-0.906
ln (Value Added)	0.877*	0.053	16.565
Technology	0.00161	0.00253	0.624
Firm Dummies: D1	0.156*	0.068	2.308
D2	0.029	0.094	0.311
D3	0.162*	0.059	2.748
Rho	0.681*	0.117	5.806

n = 66, R-squared =.985

Note: Asterisks indicate coefficients significantly different from zero.

The hypothesis of cost minimization is easily rejected by the estimates in Table 2. Recall from Equation (14) that, under cost minimization, the elasticities of derived demand sum to zero. However, the elasticity estimates in Table 2 sum to -.700, and this quantity has a t-statistic of -3.794.

We turn instead to our alternative model of constrained utility maximization. Recall from the discussion around Equation (13) that we expect all three cross-price effects to be negative. Therefore, the significant positive effect of the price of capital is *prima facie* evidence against our behavioral model. However, a positive estimate of $e(L_e, P_k)$ does not necessarily imply that the linear combinations T_1, T_2, and T_3 in Equations (16) through (18) are significantly different from zero. In particular, the positive sign on $e(L_e, P_k)$ in Equation (17) will be dampened if the share of engineering labor is sufficiently small. The only way to resolve this issue is to compute T_1, T_2, and T_3 directly.

Table 3 reports the estimates of T_1, T_2, and T_3, along with their joint significance levels, for each firm in the sample. The joint hypothesis that $T_1 = T_2 = T_3 = 0$ cannot be rejected for any of the four firms in the sample. Despite the anomalous sign on the price of capital, the tests reported in Table 3 reveal that the data are consistent with the behavioral model developed in this paper.

Table 3. Results of Hypothesis Tests, by Firm

	Restriction 1		Restriction 2		Restriction 3		Joint Significance Level	
Firm	T	Z	T	Z	T	Z	Lower bound	Upper bound
A	-0.062	-0.353	-0.098	-2.990	-0.109	-1.290	0.724	0.924
B	-0.049	-0.412	-0.128	-2.577	-0.185	-1.719	0.680	0.776
C	-0.108	-0.576	-0.118	-2.838	-0.060	-0.998	0.565	0.888
D	-0.082	-0.571	-0.126	-2.644	-0.136	-1.563	0.568	0.694

Note: T = cross-product, Z = cross-product divided by standard error.

VII. CONCLUSIONS

This paper has reported on the development and testing of a model in which firms hoard engineering labor, subject to a constraint on their profit margins imposed by the Federal Acquisition Regulations. This model implies certain restrictions on the firm's demand curve for engineering labor. These restrictions differ dramatically from those of the conventional model of cost minimization, permitting discrimination between the two models.

Empirical testing was conducted using data from four large aerospace manufacturers. Although there is some ambiguity, the data generally support the restrictions implied by our model. Hence we have evidence both that firms hoard engineering labor, and that their profit margins are effectively constrained by the Federal Acquisition Regulations.

Although hoarding of labor may appear inefficient from a short-run perspective, it may indeed be efficient from a long-run perspective. Temporary declines in business are quite common in the defense industry. Moreover, it can be quite expensive to lay off and subsequently rehire skilled engineers.

The model developed in this paper relates the demand for engineers to current market conditions only. In order to test whether apparent hoarding is optimal with respect to forecasts of future market conditions, one would need to develop a fully dynamic model of the firm. This interesting task is left for future research.

REFERENCES

[1] Department of Defense. Federal Acquisition Regulations (FAR): Department of Defense, General Services Administration, and National Aeronautics and Space Administration. Reprinted by Commercial Clearing House, Inc., Chicago, Illinois, April 1984.

[2] Averch, Harvey, and Leland Johnson. "Behavior of the Firm Under Regulatory Constraint." *American Economic Review*, vol. 52, no. 5 (December 1962), pp. 1053-1069.

[3] Williamson, Oliver. *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm.* Englewood Cliffs, N.J.: Prentice-Hall, 1964.

[4] Edwards, Franklin R. "Managerial Objectives in Regulated Industries: Expense-Preference Behavior in Banking." *Journal of Political Economy*, vol. 85, no. 1 (February 1977), pp. 147-162.

[5] Hannan, Timothy. "Expense-Preference Behavior in Banking: A Reexamination." *Journal of Political Economy*, vol. 87, no. 4 (August 1979), pp. 891-895.

[6] Osband, Kent. "Target Profitability Under DoD Profit Policy." The RAND Corporation, WD-4282-PA&E, February 1989.

[7] Rogerson, William P. "An Economic Framework for Analyzing DoD Profit Policy." Manuscript, Northwestern University, April 1988.

[8] Myers, Myron G., Donna J. S. Peterson, Paul R. McClenon, Michael J. Konvalinka, and Robert K. Wood. "Facilities Capital as a Factor in Contract Pricing." Logistics Management Institute, May 1985.

[9] Miller, Roger LeRoy. "The Reserve Labor Hypothesis: Some Tests of its Implications." *Economic Journal*, vol. 81 (March 1971), pp. 17-35.

[10] Chambers, Robert G. *Applied Production Analysis: A Dual Approach.* Cambridge: Cambridge University Press, 1988.

[11] Spann, Robert. "Rate of Return Regulation and Efficiency in Production: An Empirical Test of the Averch-Johnson Thesis." *Bell Journal of Economics and Management Science*, vol. 5, no. 1 (Spring 1974), pp. 38-52.

[12] Courville, L. "Regulation and Efficiency in the Electric Utility Industry." Bell *Journal of Economics and Management Science*, vol. 5, no. 1 (Spring 1974), pp. 53-74.

[13] Peterson, H. C. "An Empirical Test of Regulatory Effects." "Managerial Discretion and Business Behavior." *Bell Journal of Economics and Management Science*, vol. 6, no. 1 (Spring 1975), pp. 111-126.

[14] Boyes, William. "An Empirical Examination of the Averch-Johnson Effect." *Economic Inquiry*, vol. 14 (March 1976), pp. 25-35.

[15] Cowing, Thomas G. "The Effectiveness of Rate-of-Return Regulation: An Empirical Test Using Profit Functions." In *Production Economics: A Dual Approach to Theory and Applications.* Amsterdam: North-Holland, 1978.

[16] Bear, Donald V. T. "Inferior Inputs and the Theory of the Firm." *Journal of Political Economy,* vol. 73 (June 1965), pp. 287-289.

[17] Rhoades, Stephen. "Monopoly and Expense Preference Behavior: An Empirical Investigation of a Behavioralist Hypothesis." *Southern Economic Journal,* vol. 47, no. 2 (October 1980), pp. 419-432.

[18] Hannan, Timothy, and Ferdinand Mavinga. "Expense Preference and Managerial Control: The Case of the Banking Firm." *Bell Journal of Economics and Management Science,* vol. 11, no. 2 (Autumn 1980), pp. 671-682.

[19] Smirlock, Michael, and William Marshall. "Monopoly Power and Expense-Preference Behavior: Theory and Evidence to the Contrary." *Bell Journal of Economics and Management Science,* vol. 14, no. 1 (Spring 1983), pp. 166-178.

[20] Awh, Robert, and Walter Primeaux. "Managerial Discretion and Expense Preference Behavior." *Review of Economics and Statistics,* vol. 67, no. 2 (May 1985), pp. 224-231.

[21] Bailey, Elizabeth E., and John C. Malone. "Resource Allocation and the Regulated Firm." *Bell Journal of Economics and Management Science,* vol. 1, no. 1 (Spring 1970), pp. 129-142.

[22] McNicol, David L. "The Comparative Statics Properties of the Theory of the Regulated Firm." *Bell Journal of Economics and Management Science,* vol. 4, no. 2 (Autumn 1973), pp. 428-453.

[23] Crew, Michael, and Paul Kleindorfer. "Managerial Discretion and Public Utility Regulation." *Southern Economic Journal,* vol. 46 (January 1979), pp. 696-709.

[24] Arzac, Enrique, and Franklin Edwards. "Efficiency in Regulated and Unregulated Firms: An Iconoclastic View of the Averch-Johnson Thesis." In Crew, Michael, Problems in Public Utility Economics and Regulation. Lexington, Mass.: Lexington Books, 1979.

[25] "Economic Report of the President" Washington, D.C.: U.S. Government Printing Office, 1989.

[26] Evans, David S., and James J. Heckman. "Multiproduct Cost Function Estimates and Natural Monopoly Tests for the Bell System." In *Breaking Up Bell.* Amsterdam: North-Holland, 1982.

[27] Hausman, Jerry A. "Specification Tests in Econometrics." Econometrica, vol. 46, no. 6 (November 1978), pp. 1251-1271.

[28] Harmon, Bruce R., J. Richard Nelson, Mitchell S. Robinson, Kathryn L. Wilson, and Steven R. Shyman. "Military Tactical Aircraft Development Costs." 5 volumes. Institute for Defense Analyses, R-339, September 1988.

APPENDIX A.

HYPOTHESIS TESTING

APPENDIX A: HYPOTHESIS TESTING

SAMPLING DISTRIBUTION

Our behavioral model implies three linear restrictions, which are reproduced here:

(A-1) $T_1 = C_p e(L_e, P_e) - C_e e(L_e, P_p) = 0$,

(A-2) $T_2 = C_k e(L_e, P_e) - C_e e(L_e, P_k) = 0$,

(A-3) $T_3 = C_m e(L_e, P_e) - C_e e(L_e, P_m) = 0$.

In these equations, $e(L_e, P_j)$ is the elasticity of demand for engineering labor with respect to the j^{th} input price, and C_j is expenditure on the j^{th} input divided by total expenditure on all inputs.

The elasticities in Equations (A-1) through (A-3) correspond to the regression coefficients in a log-linear specification. These regression coefficients are estimated from finite samples, and are subject to sampling variation. The magnitude of the sampling variation is expressed by the standard errors, which are available from the regression output.

We estimated the cost shares for each firm as the simple averages over the years observed in the sample. Strictly speaking, sampling variation in the average shares introduces additional randomness into T_1 through T_3, beyond that embodied in the regression coefficients. We choose to ignore this variation, instead treating the shares for each firm as fixed constants. In so doing, we underestimate the total variance in T_1 through T_3, making it more likely that we will reject the proportionality restrictions. In effect, we bias the test against our model; if the proportionality restrictions hold despite this bias, then we can have additional confidence in their validity.

If the regression disturbances are normally distributed, then the regression coefficients (i.e., the elasticity estimates) are themselves normally distributed. If we treat the cost shares as fixed constants, then the test statistics T_1 through T_3, being linear combinations of normally-distributed regression coefficients, are again normally distributed.

If our behavioral model is valid, then the test statistics each have a mean of zero. Of course, the computed values in any finite sample might still differ from zero, due to sampling variation. To determine whether the departure from zero is statistically significant, we require estimates of the standard errors of T_1 through T_3.

Again treating the cost shares as fixed constants, the standard errors of T_1 through T_3 may be obtained from the elementary formula for the variance of a linear combination. Letting $e_e = e(L_e, P_e)$, $e_j = e(L_e, p_j)$, and $T_j = C_j e_e - C_e e_j$:

(A-4) $\quad Var(T_j) = C_j^2 Var(e_e) + C_e^2 Var(e_j) - 2C_j C_e Cov(e_e, e_j)$.

If $Var(T_j)$ were known with certainty, then the ratio of T_j to the square root of $Var(T_j)$ would have a standard normal distribution. In this case, tests of significance could be made with reference to tables of the normal distribution.

Unfortunately, the terms in Equation (A-4) are not known with certainty, but are instead estimated from finite samples along with the regression coefficients. In this situation, tests of significance are based on the ratio of T_j to the square root of the *sample estimate* of $Var(T_j)$. It can be shown that this ratio has a t-distribution.[1] The t-distribution has "fatter tails" than the standard normal distribution, so that hypothesis tests are somewhat less precise and confidence intervals are somewhat wider when based on the t-distribution.

When testing the hypothesis $T_j = 0$, we must decide upon either a one-sided or two-sided alternative. The one-sided alternative would be appropriate if we believed that failure of the hypothesis was most likely in a single direction (e.g., if T_j is not zero, then most likely $T_j < 0$). Because we held no such beliefs, we tested all hypotheses using two-sided alternatives. Therefore, we computed the significance level as the probability, under the null hypothesis, of obtaining a test statistic larger *in absolute value* than that estimated from the sample:

(A-5) $\quad p = Pr(|T| > |T^*|$ given hypothesis is true).

In Equation (A-5), T is a random variable drawn from the t-distribution, and T* is the test statistic estimated from the sample.

1 See Schmidt ([A-1], p. 19). The t-distribution converges to the standard normal distribution as the sample size approaches infinity. However, for the samples sizes in this paper, the numerical differences between the two distributions are still significant.

Finally, recall that we are testing the *joint* hypothesis that $T_1 = T_2 = T_3 = 0$ for each firm. Hence we must combine the individual significance levels for T_1 through T_3 into an overall significance level. This problem of multiple comparisons is discussed in the follow subsection.

MULTIPLE COMPARISONS

The problem of multiple comparisons often arises in statistical applications. For example, consider testing a sample of individuals suffering from hypertension, to determine the efficacy of a new drug. The null hypothesis is that blood pressure for the treatment group is no different from blood pressure for a control group, where the latter group may be given a placebo.

If a single measurement of blood pressure is taken for each individual, then hypothesis testing may proceed in the usual fashion, using any desired significance level. For example, suppose that the five-percent significance level is chosen. The hypothesis test might be based on the difference between average blood pressure in the two groups. The null hypothesis is rejected if and only if this difference is so large that it would occur with probability five percent or less, *if the null hypothesis were true.*

Now suppose that the two groups are stratified on the basis of some other characteristics, such as age, marital status, and whether or not the individual smokes cigarettes. The stratification scheme defines cells, and each individual falls into exactly one such cell. Average blood pressure may now be compared for the treatment and control groups, *within each cell* (e.g., compare young, married smokers in the treatment group to young, married smokers in the control group).

Suppose there are 20 cells, and suppose that the null hypothesis of equal blood pressure is true. Recall that we allow a five-percent chance of rejecting any null hypothesis when that hypothesis is true. Hence, among our 20 comparisons of blood pressure, we expect exactly one (20 times .05) significant difference to arise *simply by chance.* It follows that the significance level of the entire sequence of 20 tests is much higher than the nominal significance level of five percent.

A similar problem arises when multiple measurements of blood pressure are taken for each individual. Multiple measurements may be taken, for example, in order to provide early detection of undesired side effects of the drug. If the null hypothesis of equal blood pressure is true, there is still a five-percent chance of finding "significant" differences on any testing date. With multiple testing dates, the probability of rejecting the null hypothesis is obviously much greater than five percent.

The exact significance level of the sequence of tests may be computed from the individual significance levels, if we know the degree of dependence between the tests. At one extreme, suppose that the 20 tests for each individual were conducted one immediately after the other, without interruption. Then the 20 tests would yield essentially identical readings and, in effect, there would be only one reading per individual. In this case, the overall significance level would equal the nominal significance level of five percent.

As an intermediate case, suppose the $n = 20$ tests for each individual were statistically independent. The probability of rejecting the null hypothesis on any date is given by $p = .05$. Under independence, the number of dates on which rejection occurs has a binomial distribution with parameters (n, p). The null hypothesis is rejected overall if it is rejected *on any date (or dates).* The number of dates on which rejection occurs has expected value $np = 1.0$, so we reject the overall hypothesis "on average." More precisely, the probability of overall rejection is the complement of the probability that rejection does not occur at any date:

(A-6) $$p^* = 1 - (1-p)^n = 1 - .95^{20} = .6415.$$

Equation (A-6) may be generalized to the case in which the probabilities of rejection on each date are not equal. If p_t is the probability of rejection on date t, the generalization of Equation (A-6) is:

(A-7) $$p^* = 1 - \prod (1 - p_t) \ .$$

Finally, we explore the other extreme case in order to develop an upper bound on p^*. Consider the case in which rejection on any one date precludes rejection on any other date, so that rejection can occur on at most one date but not on multiple dates. In this case, the overall probability of rejection is the sum of the mutually exclusive probabilities on each date:

(A-8) $$p^* = \Sigma p_t \ .$$

Equation (A-8) yields an upper bound on p^*, and the addition of an appropriate lower bound yields the so-called Bonferroni inequality[2]:

(A-9) $$\text{Max } p_t \leq p^* \leq \Sigma\, p_t \ .$$

---

2 See Barlow and Proschan ([A-2], p.25) or Feller ([A-3], p. 110). The Bonferroni inequality is also known as the "inclusion-exclusion method."

The Bonferroni bound is particularly tight when one of the p_t is much larger than the others. For example, suppose that $p_1 = Max(p_t) = .10$ and $p_2 = p_3 = .01$. Then the Bonferroni bound is $.10 \leq p^* \leq .12$. The lower bound would be exact if rejection on date $t = 2$ or $t = 3$ guaranteed rejection on date $t = 1$, and the upper bound would be exact if rejection on one date precluded rejection on any other date (i.e., the mutually exclusive case). By contrast, if the three dates were independent, Equation (A-7) would yield the exact probability of $p^* =.1179$.

For the applications in this paper, the sequence of tests was be conducted using overlapping sets of regression coefficients, computed from a common data sample. Because the independence assumption clearly fails, naive application of equation (A-7) would yield spurious results. Instead, we applied the Bonferroni bounds given in (A-9). Fortuitously, the Bonferroni bounds turned out to be quite tight.

REFERENCES

[A-1] Schmidt, Peter. *Econometrics*. New York: Marcel Dekker, 1976.

[A-2] Barlow, Richard, and Frank Proschan. *Statistical Theory of Reliability and Life Testing*. New York: Holt, Reinhart and Winston, 1981.

[A-3] Feller, William. *An Introduction to Probability Theory and its Applications, Volume I*. New York: Wiley, 1968.