nation of Cobb-Douglas
959).
ιips Using Instrumental

ιations and the Cobb-
7, Faculty of Commerce
ingham, England, June

7

Returns to Scale in Electricity Supply

MARC NERLOVE, *Stanford University*

The study of returns to scale in public-utility enterprises has a long, if not always honorable, history. The question of whether there are increasing or decreasing returns to scale and over what range of output has, as we know, an important bearing on the institutional arrangements necessary to secure an optimal allocation of resources. If as many writers in the field appear to believe, there are increasing returns to scale over the relevant range of outputs produced by utility undertakings, then these companies must either receive subsidies or resort to price discrimination in order to cover costs at socially optimal outputs.

In addition, as Chenery [2] has pointed out, the extent of returns to scale is a determinant of investment policies in growing industries. If there are increasing returns to scale and a growing demand, firms may find it profitable to add more capacity than they expect to use in the immediate future.

In studying the problem of returns to scale, the first question one must ask is "To what use are the results to be put?" It is inevitable that the purpose of an analysis should affect its form. In particular, the reason for obtaining an estimate of returns to scale will affect the *level* of the analysis: industry, firm, or plant. For many questions of pricing policy, for example, the plant is the relevant entity. On the other hand, when questions of taxation are at issue, the industry may be the appropriate unit of analysis. But if we are concerned primarily with the general question of public regulation and with investment decisions and the like, it would seem that the economically relevant entity is the firm. Firms, not plants are regulated, and it is at the level of the firm that investment decisions are made.

The U.S. electric power industry is a regulated public utility. Privately

I am indebted for a great deal of helpful advice to I. Adelman, K. J. Arrow, A. R. Ferguson, W. R. Hughes, S. H. Nerlove, P. A. Samuelson, and H. Uzawa. Had I been able to take all the advice I received, perhaps I could lay a part of the blame for the deficiencies of this paper on these people. The situation, however, is otherwise

Support, in part, for the research on which this paper is based has been received from the Department of Economics, University of Minnesota, and under a grant from the Rockefeller Foundation at Stanford University. Stenographic assistance was received from the Office of Naval Research under Contract Nonr-225(50).

I wish also to acknowledge the help of M. S. Arora, G. Fishman, J. Johnston, H. Kanemitsu, and N. K. Rao, who performed the computations on which this paper is based.

owned firms, with which I am exclusively concerned in this study, account for nearly 80 per cent of all power produced. The technological and institutional characteristics of the electric power industry that are important for the model I shall develop are as follows:

1. Power cannot be economically stored in large quantities and, with few exceptions, must be supplied on demand.
2. Revenues from the sale of power by private companies depend primarily on rates set by utility commissions and other regulatory bodies.
3. Much of the fuel used in power production is purchased under long-term contracts at set prices. The level of prices is determined in competition with other uses.
4. The industry is heavily unionized, and wage rates are also set by contracts that extend over a period of time. Over long periods, wages appear to be determined competitively.
5. The capital market in which utilities seek funds for expansion is highly competitive and the rates at which individual utilities can borrow funds are little affected by individual actions over a wide range. Construction costs vary geographically and also appear to be unaffected by an individual utility's actions.

From these characteristics we may draw two conclusions, which lead to the model presented below. First, it is plausible to regard the output of a firm and the prices it pays for factors of production as exogenous, despite the fact that the industry does not operate in perfectly competitive markets. Second, the problem of the individual firm in the industry would appear to be that of minimizing the total costs of production of a given output, subject to the production function and the prices it must pay for factors of production. I shall adopt this last conclusion in what follows, although it is subject to some qualifications.

There are two basic objections to the cost-minimization hypothesis. First, rates in the industry are governed by a "cost plus" principle designed to secure investors "a fair return on fair value" (whatever that may mean). Although the application of this principle is a complicated matter in practice, it is clear that if a utility minimizes costs too much, i.e., decreases its costs to such an extent that, under the current rate structure, it obtains a substantial increment in earnings, the regulatory body may initiate an investigation and wipe out the increment through a decrease in rates. My impression, however, is that most utilities operate at a considerable distance from this "danger point."

A second objection to the cost-minimization hypothesis is that it is implicitly static; i.e., it does not reflect the fact that utilities are less concerned with cost minimization at a *point in time* than they are with minimization *over time*. In a dynamic formulation capital costs may be particularly affected. However, two contrary tendencies seem to exist: On the one hand, a steady rate of technological improvement has been experienced and may be expected to continue in this industry; thus, it is advantageous to postpone investment commitments. On the other hand, if there are increasing returns to scale, the steady growth in demand might be expected, *à la* Chenery [2], to lead to capital expenditures in excess of current needs. This tendency to over-capitalization may be aided and abetted by rate commissions, which are often inclined to support it after the fact through an increase in rates.

A related objection has been raised by William Hughes. He pointed out, in effect, that the existence of several power pools among companies treated separately in my analysis means that the outputs of such companies may not be truly exogenous as I have assumed.

Previous empirical investigations that have a bearing on returns to scale in electricity supply are those of Johnston [10, pp. 44-73], Komiya [11], Lomax [12], and Nordin [16]. All of these are concerned with returns to scale at the level of the plant, not the firm, and present evidence which suggests that there are increasing or constant returns to scale in the production of electricity. It is shown in Appendix A, however, that because of transmission losses and the expenses of maintaining and operating an extensive transmission network, a firm may operate a number of plants at outputs in the range of increasing returns to scale and yet be in the region of decreasing returns when considered as a unit. Although firms as a whole have been treated in this investigation, the problem of transmission and its effects on returns to scale has not been incorporated in the analysis, which relates only to the *production* of electricity. The results of this analysis are in agreement with those of previous investigators and suggest that the bulk of privately owned U.S. utilities operate in the region of increasing returns to scale, as is generally believed. Nevertheless, the results also suggest that the *extent* of returns to scale at the firm level is overestimated by analyses that deal with individual plants.

As indicated in Table 1, the production of electric power is carried out in three main ways:

1. By internal combustion engines. This method accounts for a negligible fraction of the power produced.
2. By hydroelectric installations. This method accounts for about one-third of all U.S. power production.
3. By steam-driven installations. This method accounts for the remaining two-thirds of U.S. power production.

Few firms rely solely on hydroelectric production because of the unreliability of supply. Furthermore, suitable sites for hydroelectric installations are rather limited and, except for those sites requiring an immense capital investment, almost fully exploited. Because of the great qualitative difference between steam and hydraulic production of electricity, this analysis is limited to steam generation. Since the variable costs of hydroelectric production are extremely low and it appears that firms fully exploit these possibilities, neglect of hydraulic generation should little affect the results on returns to scale.

The costs of steam-electric generation consist of (a) energy costs, and (b) capacity costs. The former consist mainly of the costs of fuel, of which coal is the principal one (see Table 2). Energy costs tend to vary with total output, and depend little on the distribution of demand through time. Capacity costs include interest, depreciation, maintenance, and most labor costs; these costs tend to vary, not with total output, but with the maximum anticipated demand for power (i.e., the peak load). Unfortunately, available data do not permit an adequate treatment of the peak-load dimension of output, hence it has been neglected in this study.

Even if the temporal distribution of demand does not differ systematically from one size firm to another, however, the results may be affected. A large firm with many plants and operating over a wide area has a greater

TABLE 1

PER CENT OF TOTAL KILOWATT-HOURS PRODUCED BY TYPE OF PLANT, 1930–1950, U.S.

Year	Steam Generating Plants	Hydroelectric Installations	Internal Combustion Engines
1930	65.1	34.2	0.7
1940	65.6	33.4	1.0
1950	69.8	29.1	1.1

TABLE 2

PER CENT OF TOTAL STEAM-ELECTRIC GENERATION (KWH) BY TYPE OF FUEL, 1930–1950, U.S.

Year	Coal	Oil	Gas
1930	84.8	4.7	10.5
1940	81.9	6.6	11.5
1950	66.4	14.5	19.1

Source: R. E. Caywood, *Electric Utility Rate Economics*. New York: McGraw-Hill, 1956.

diversity of customers; hence, a large firm is more likely to have a peak load that is a small percentage of output than a small firm. It follows that capacity costs per unit of output tend to be less for larger firms. But this is a real economy of scale, and one reason for looking at firms rather than plants is precisely to take account of such phenomena. Of course, explicit introduction of peak-load characteristics would be better than the implicit account that is taken here.

1. The Model Used

As indicated, the characteristics of the electric power industry suggest that a plausible model of behavior is cost minimization, and that output and factor prices may be treated as exogenous. This suggests that traditional estimation of a production function from cross-section data on inputs and output is incorrect; fortunately, it also suggests a correct procedure. Let

c = total production costs,

y = output (measured in kwh),

x_1 = labor input,

x_2 = capital input,

x_3 = fuel input,

p_1 = wage rate,

p_2 = "price" of capital,

p_3 = price of fuel,

u = a residual expressing neutral variations in efficiency among firms.

Suppose that firms have production functions of a generalized Cobb-Douglas type:

$$(1) \qquad y = a_0 x_1^{a_1} x_2^{a_2} x_3^{a_3} u \,.$$

Minimization of costs,

$$(2) \qquad c = p_1 x_1 + p_2 x_2 + p_3 x_3 \,,$$

implies the familiar marginal productivity conditions:

$$(3) \qquad \frac{p_1 x_1}{a_1} = \frac{p_2 x_2}{a_2} = \frac{p_3 x_3}{a_3} \,.$$

If the efficiency of firms varies neutrally,[1] as indicated by the error term in (1), and the prices paid for factors vary from firm to firm, then the levels of input are not determined independently but are determined jointly by the firm's efficiency, level of output, and the factor prices it must pay. In short, a fitted relationship between inputs and output is a *confluent* relation that does not describe the production function at all but only the net effects of differences among firms. (For a more general discussion, see [13, 15].)

In such cases, however, it may be possible to fit the *reduced form* of a system of structural relations such as (1) and (3) and to derive estimates of the structural parameters from estimates of the reduced-form parameters. Not only does it turn out to be possible in this case, but an important reduced form turns out to be the cost function

$$c = k y^{1/r} p_1^{a_1/r} p_2^{a_2/r} p_3^{a_3/r} v \,, \tag{4}$$

where

$$k = r(a_0 a_1^{a_1} a_2^{a_2} a_3^{a_3})^{-1/r} \,,$$

$$v = u^{-1/r} \,,$$

and

$$r = a_1 + a_2 + a_3 \,.$$

The parameter r measures the degree of returns to scale. The fundamental duality between cost and production functions, demonstrated by Shephard [17], assures us that the relation between the cost function, obtained empirically, and the underlying production function is unique.[2] Under the cost minimization assumption, they are simply two different, but equivalent ways of looking at the same thing.

Note that the cost function must include factor prices if the correspondence is to be unique. The problem of changing (over time) or differing (in a cross section) factor prices is an old one in statistical cost analysis; see [10, pp. 170-76]. Most generally, it seems to have been handled by deflating cost figures by an index of factor prices, a procedure that Johnston [10] shows typically leads to bias in the estimation of the cost

[1] A model incorporating non-neutral variations in efficiency of the form

$$y = (a_0 u_0) x_1^{a_1 u_1} x_2^{a_2 u_2} x_1^{a_3 u_3}$$

was discussed in my paper "On Measurement of Relative Economic Efficiency," abstract, *Econometrica*, **28** (July 1960), 695. It is interesting to note that despite the complex way in which the random elements u_0, u_1, and u_2 enter, there are circumstances under which it is possible to estimate the parameters in such a production function.

[2] I owe this point to Hirofumi Uzawa. It is true, of course, only if all firms have the same production function, except perhaps for differences in the constant term, so that aggregation difficulties may be neglected.

curve unless correct weights, which depend on (unknown) parameters of the production function, are used. It seems strange that no one has taken the obvious step of *including factor prices directly in the cost function.* If price data are available for the construction of an index and prices do not move proportionately, in which case no bias would result from deflation, why not use the extra information afforded?

What form of production function is appropriate for electric power? The generalized Cobb-Douglas function presented above is attractive for two reasons: First, it leads to a cost function that is linear in the logarithms of the variables

$$C = K + \frac{1}{r}Y + \frac{a_1}{r}P_1 + \frac{a_2}{r}P_2 + \frac{a_3}{r}P_3 + V\,, \tag{5}$$

where capital letters denote logarithms of the corresponding lower-case letters. The linearity of (5) makes it especially easy to estimate. Second, a single estimate of returns to scale is possible (it is the reciprocal of the coefficient of the logarithm of output), and returns to scale do not depend on output or factor prices. (The last-mentioned advantage turns out to be a defect as we shall see when we come to examine a few statistical results.) But does such a function accurately characterize the conditions of production in the electric power industry?

A casual examination of trade publications suggests that once a plant is built, fixed proportions are more nearly the rule. Support for this view is given by Komiya [11], who found that data on inputs and output for individual plants were better approximated by a fixed-proportions model that allowed differences in the proportions due to scale. A simplified version of Komiya's model is[3]

$$\begin{aligned} x_1 &= a_1 y^{b_1}\,, \\ x_2 &= a_2 y^{b_2}\,, \\ x_3 &= a_3 y^{b_3}\,. \end{aligned} \tag{6}$$

At the firm level, however, there are many possibilities for substitution that may go unnoticed at the plant level; for example, labor and fuel may be substituted for capital by using older, less efficient plants more intensively or by using a large number of small plants rather than a few large ones.

[3] Since y is exogenous, it would be appropriate to estimate the coefficients in (6) by least squares. An objection to this, however, is the fact that, if individual plants are considered, the output allocated to *each* is not exogenous; see Westfield [19, pp. 15-81). Furthermore, Komiya does not use output but name-plate rated capacity and input levels adjusted to full capacity operation. It is even more doubtful whether the former can be considered as exogenous in a cross section. My objection here is closely related to the one raised by Hughes (see p. 169); however, while the endogenicity of output at the plant level is clear, its endogenicity at the firm level for a member of a power pool is conjectural

Given persistent differences in the factor prices paid by different firms, cross-section data should reflect such possibilities of substitution. Certainly, as a provisional hypothesis, a generalized Cobb-Douglas function may be appropriate.

It would, of course, be preferable to *test* whether significant substitution among factors occurs at the firm level. The use of the generalized Cobb-Douglas unfortunately does not permit us to do so except in a very general way, since its form implies that the elasticity of substitution between any pair of factors is one. A more general form, which has both the Cobb-Douglas and fixed coefficients as limiting cases, has recently been suggested by Arrow, Minhas, Chenery, and Solow [1]. Constant returns to scale are assumed, but the form can be easily generalized; in a more general form it is

$$y = [a_1x_1^b + a_2x_2^b + a_3x_3^b]^{1/f}\,. \tag{7}$$

In this case returns to scale are given by the ratio b/f and the elasticity of substitution between any pair of factors can be shown to be $1/(1-b)$. In the special case in which $b = f$ it can be shown that the limiting form of (7) as the elasticity of substitution goes to zero is

$$y = \min\left\{\frac{x_1}{(a_1+a_2+a_3)^{1/b}-1},\frac{x_2}{(a_1+a_2+a_3)^{1/b}-1},\frac{x_3}{(a_1+a_2+a_3)^{1/b}-1}\right\}, \tag{8}$$

or fixed coefficients, and the limiting form as the elasticity of substitution goes to one is

$$y = (a_1+a_2+a_3)^{1/b}x_1^{a_1/(a_1+a_2+a_3)}x_2^{a_2/(a_1+a_2+a_3)}x_3^{a_3/(a_1+a_2+a_3)}\,, \tag{9}$$

or Cobb-Douglas. Although I have not formally demonstrated the fact, it is possible that the limiting form of the more general case (7) is something like the Komiya model as the elasticity of substitution tends to zero, and like the generalized Cobb-Douglas as it tends to one.

Unfortunately, in its generalized form (7) is quite difficult to estimate from the data available. Furthermore, although clearly superior to the generalized Cobb-Douglas form, (7) still implies that the elasticity of substitution between any pair of factors (e.g., labor capital and fuel capital) is the same, which hardly seems reasonable. Other generalizations are possible, but none that I have found thus far offers much hope of being amenable to a reasonable estimation procedure.

If the generalized Cobb-Douglas form is adopted, however, relatively simple estimation procedures can be devised for evaluating the parameters of the production function. The reduced form of (1) and (3) that incorporates all but one of the restrictions on the parameters in the derived demand equations (which are the more usual reduced form) is nothing but the cost function.

The only restriction not incorporated in (4) or (5) is that the coefficients of the prices must add up to one. It is a simple matter to incorporate this restriction, however, by dividing costs and two of the prices by the remaining price (it doesn't matter either economically or statistically which price we choose). When fuel price is used as the divisor, the result is

$$C - P_3 = K + \frac{1}{r} Y + \frac{a_1}{r}(P_1 - P_3) + \frac{a_2}{r}(P_2 - P_3) + V \,, \tag{10}$$

which will be called Model A.

Model A assumes that we have relevant data on the "price" of capital and that this price varies significantly from firm to firm. If neither is the case, we are in trouble. Most of the results presented here are based on Model A, but the data used for this price of capital are clearly inadequate. (See Appendix B.) If one supposes, however, that the price of capital is the same for all firms, which is not implausible, one can do without data on capital price and use the restriction on the coefficients of output and prices to estimate the elasticity of output with respect to capital input. The assumption that capital price is the same for all firms implies

$$C = K' + \frac{1}{r} Y + \frac{a_1}{r} P_1 + \frac{a_3}{r} P_3 + V \,, \tag{11}$$

where $K' = K + (a_2/r)P_2$, since the exponents of the input levels in (1) are assumed to be the same for all firms. Equation (11) is called Model B.

2. Some Statistical Results and Their Interpretation

Estimation of Model A from a cross section of firms requires that we obtain data on production costs, total physical output, and the prices of labor, capital, and fuel for each firm; for Model B we do not need the price of capital, since it is assumed to be the same for all firms. Details of the construction of these data for a sample of 145 privately owned utilities in 1955 are given in Appendix B and are not discussed here at any length. Suffice it to say that these data are far from adequate for the purpose, and I now believe that a better job could have been done with other sources.

The results from the least-squares regression suggested by equation (10) are given in line I of Table 3; the interpretation of these results in terms of the parameters of the production function is given in line I of Table 4. The R^2 is 0.93, which is somewhat unusual for such a large number of observations; increasing returns to scale are indicated, and the elasticities of output with respect to labor and fuel have the right sign and are of plausible magnitude; however, the elasticity of output with respect to capital price has the wrong sign (fortunately, it is statistically insignificant).

TABLE 3

RESULTS FROM REGRESSIONS BASED ON MODEL A FOR 145 FIRMS IN 1955

Regression No.	Coefficient				R^2
	Y	$P_1 - P_3$	$P_2 - P_3$	x	
I	0.721 (±.175)	0.562 (±.198)	−0.003 (±.192)	—	0.931
II	0.696 (±.173)	0.512 (±.199)	0.033 (±.185)	−0.045 (±.022)	0.932
IIIA	0.398 (±.079)	0.641 (±.691)	−0.093 (±.669)	—	0.512
IIIB	0.668 (±.116)	0.105 (±.275)	0.364 (±.277)	—	0.635
IIIC	0.931 (±.198)	0.408 (±.199)	0.249 (±.189)	—	0.571
IIID	0.915 (±.108)	0.472 (±.174)	0.133 (±.157)	—	0.871
IIIE	1.045 (±.065)	0.604 (±.197)	−0.295 (±.175)	—	0.920
IVA	0.394 (±.055)			—	
IVB	0.651 (±.189)			—	
IVC	0.877 (±.376)	0.435 (±.207)	0.100 (±.196)	—	0.950
IVD	0.908 (±.354)			—	
IVE	1.062 (±.169)			—	

Figures in parentheses are the standard errors of the coefficients.

The dependent variable in all analyses was $C - P_3$.

The variables are defined as follows:

C = log costs Y = log output P_1 = log wage rate P_2 = log capital "price"

$$P_3 = \text{log fuel price} \qquad x = \frac{\text{output 1955} - \text{output 1954}}{\text{output 1954}}$$

RETURNS TO SCALE A
FROM

Regression No.	I
I	
II	
IIIA	
IIIB	
IIIC	
IIID	
IIIE	
IVA	
IVB	
IVC	
IVD	
IVE	

The difficultie
tered in measuri
were measured
long-term debt
price was comp
index of constr
purchases of cap
it does not; it i
effect on costs.
on Model B are
regression for th
of Table 6. It i
elasticities of ou

TABLE 4

RETURNS TO SCALE AND ELASTICITIES OF OUTPUT WITH RESPECT TO VARIOUS INPUTS DERIVED FROM RESULTS PRESENTED IN TABLE 3 FOR 145 FIRMS IN 1955

Regression No.	Returns to Scale	Elasticity of Output with Respect to		
		Labor	Capital	Fuel
I	1.39	0.78	−0.00	0.61
II	1.44	0.74	0.01	0.69
IIIA	2.52	1.61	−0.02	0.93
IIIB	1.50	0.16	0.53	0.81
IIIC	1.08	0.44	0.27	0.37
IIID	1.09	0.52	0.15	0.42
IIIE	0.96	0.58	−0.29	0.67
IVA	2.52	1.10	0.25	1.17
IVB	1.53	0.65	0.15	0.73
IVC	1.14	0.50	0.11	0.53
IVD	1.10	0.48	0.11	0.51
IVE	0.94	0.41	0.09	0.44

The difficulties with capital may be due in part to the difficulty I encountered in measuring both capital costs and the price of capital. The former were measured as depreciation charges plus the proportion of interest on long-term debt attributable to the production plant; the figure for capital price was compounded of the yield on the firm's long-term debt and an index of construction costs. Depreciation figures reflect past prices and purchases of capital equipment, whereas the price of capital as I constructed it does not; it is perhaps not so surprising then that the price has little effect on costs. Model B is designed to evade this difficulty. Results based on Model B are presented in line V of Table 5 and the implications of this regression for the parameters in the production function are given in line V of Table 6. It is apparent that the estimates of returns to scale and the elasticities of output with respect to labor and fuel are changed very little;

TABLE 5

RESULTS FROM REGRESSIONS BASED ON MODEL B FOR 145 FIRMS IN 1955. DEPENDENT VARIABLE WAS C = LOG COSTS

Regression No.	Coefficient			R^2
	Y	P_1	P_3	
V	0.723 (±.019)	0.483 (±.303)	0.496 (±.106)	0.914
VIA	0.361 (±.086)	0.212 (±1.259)	0.655 (±.350)	0.438
VIB	0.661 (±.106)	−0.401 (±.333)	0.490 (±.134)	0.672
VIC	0.985 (±.180)	−0.014 (±.261)	0.330 (±.138)	0.647
VID	0.927 (±.106)	0.327 (±.228)	0.426 (±.064	0.884
VIE	1.035 (±.067)	0.704 (±.272)	0.643 (±.132	0.934

Figures in parentheses are the standard errors of the coefficients.

TABLE 6

RETURNS TO SCALE AND ELASTICITIES OF OUTPUT WITH RESPECT TO VARIOUS INPUTS DERIVED FROM RESULTS PRESENTED IN TABLE 5 FOR 145 FIRMS IN 1955.

Regression No.	Returns to Scale	Elasticity of Output with Respect to		
		Labor	Capital	Fuel
V	1.38	0.67	0.0[illegible]	0.69
VIA	2.77	0.59	1.3[illegible]	0.74
VIB	1.51	−0.62	0.6[illegible]	0.33
VIC	1.02	−0.01	0.2[illegible]	0.46
VID	1.08	0.35	−0.3[illegible]	0.62
VIE	0.97	0.68	0.0[illegible]	0.68

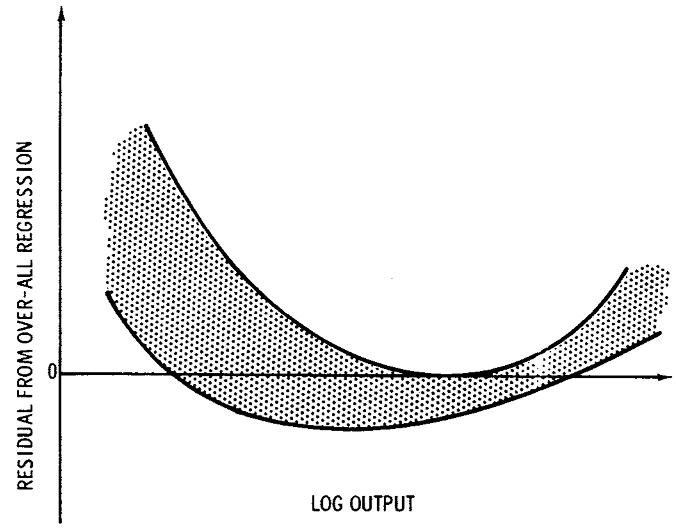

FIG. 1

the elasticity with respect to capital is of the right sign but still unreasonably low for an industry that is so capital-intensive.[4]

A second difficulty with these regressions is not apparent from an examination of the coefficients and their standard errors. As part of these analyses, the residuals from the regressions were plotted against the logarithm of output. The result is schematically pictured in Fig. 1. It is clear that neither regression relationship is truly linear in logarithms. To test this visual impression the observations were arranged in order of ascending output, and Durbin-Watson statistics were computed; the values of the statistics indicated highly significant positive serial correlation, which confirmed the visual evidence.

Aside from difficulties with the basic data, there appear to be at least two plausible and interesting hypotheses accounting for the result.

[4]K. Arrow has pointed out that considerations of plausibility implicitly involve an alternative method of estimating the coefficients in the production function: From the marginal productivity conditions (3), we find that for any pair of inputs i and j,

$$\frac{p_i x_i}{p_j x_j} = \frac{a_i}{a_j}.$$

Hence, by constructing some average of the ratios of expenditures on factors, we obtain estimates of the ratios of exponents in the production function. Had the data been arranged in such a manner as to facilitate computation of expenditures on individual factors, a comparison of the ratios a_i/a_j obtained in this way with those derived from the cost function would have been a useful supplement to the analysis. Arrow also pointed out that one could also verify the results by the fit of the production function derived from them. Unfortunately, it is not feasible to obtain good physical measures of the inputs and such measures are required for this test.

RMS IN 1955.

R^2
0.914
0.438
0.672
0.647
0.884
0.934

OUS INPUTS DERIVED
1955.

Respect to
Fuel
0.69
0.74
0.33
0.46
0.62
0.68

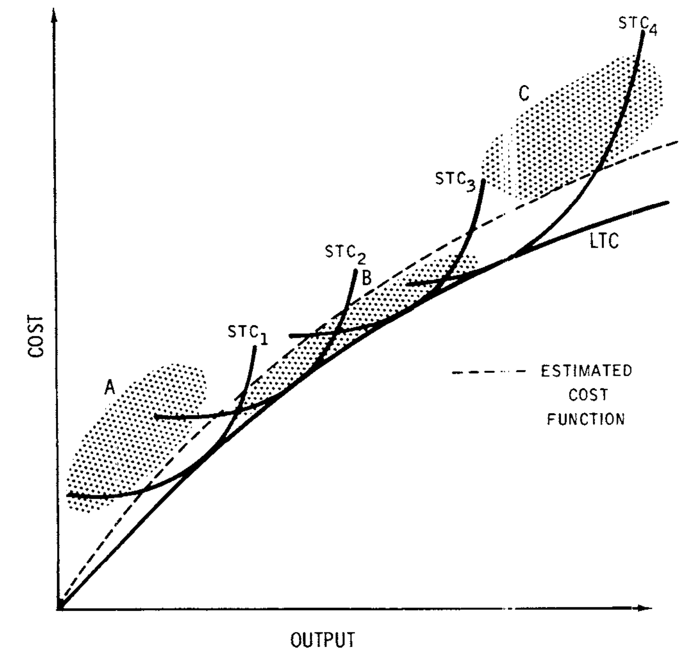

FIG. 2

1. The first explanation of the result derives from a dynamic considerations closely related to those underlying Friedman's Permanent-Income Hypothesis [7]. The important thing to note is that actual costs are underestimated by the regressions at both high and low outputs. Consider the situation pictured in Fig. 2. Firms operate not on the long-run cost curve, but at points on the various short-run curves. If firms are evenly distributed about their optimal outputs (i.e., outputs at which long-run marginal cost equals short-run marginal cost), the effect will be to increase the estimate of the extent of increasing returns to scale if they are increasing, or diminish further the estimate of returns to scale if they are decreasing.[5] But elsewhere Friedman holds that a uniform distribution is not likely to occur; in fact he says, "The firms with the largest output are unlikely to be producing at an unusually low level: on the average they are likely to be producing at an unusually high level; and conversely for those that have the lowest output" [14, p. 237].

The situation described by Friedman is pictured in Fig. 2 by the shaded areas A, B, and C, which refer, respectively, to observations on firms with unusually low, usual, and unusually high outputs. The Friedman explana-

[5] This argument rests partly on the form of the function that constrains it to pass through the origin.