

GRADUATE SCHOOL
OF INDUSTRIAL ENGINEERING
CARNegie-MELLON UNIVERSITY

EX 125
16 1972

REPRINT NO. 938

On Parsimonious Explanations of Production Relations

by

Herbert A. Simon

1979



Carnegie Mellon University
PITTSBURGH, PENNSYLVANIA 15213

Graduate School of Industrial Administration

William Larimer Mellon, Founder

Carnegie-Mellon University

Pittsburgh, Pennsylvania 15213

SOME CURRENT REPRINTS

804. Universal Planning Horizons for Generalized Convex Production Scheduling. Thomas E. Morton.
805. Behaviorally Anchored Rating Scales: Some Theoretical Issues. Robert S. Atkin and Edward J. Conlon.
806. Problem-Solving Strategies and the Writing Process. Linda S. Flower and John R. Hayes.
808. Capital Markets and the Short Run Behavior of Life Cycle Savers. Walter Dolde.
809. An Improved Algorithm for the Stationary Cost Dynamic Lot Size Model with Backlogging. Thomas E. Morton.
810. Further Reflections on Wage Controls. Comment by Arnold R. Weber and Daniel J. B. Mitchell. Reply by George P. Schultz and Kenneth W. Dam.
811. Attitude Change or Attitude Formation? An Unanswered Question. Carnegie-Mellon University Marketing Seminar.
812. Market Risk Adjustment in Project Valuation. George M. Constantinides.
813. A Least-Element Theory of Solving Linear Complementarity Problems as Linear Programs. Richard W. Cottle and Jong-Shi Pang.
814. On Cone Orderings and the Linear Complementarity Problem. Jong-Shi Pang.
815. A Search for Testable Implications of the Tiebout Hypothesis. Dennis Epple, Allan Zeleznitz, and Michael Visscher.
817. Optimal Investment, Pricing and Replacement of Computer Resources. Charles H. Kriebel, Anthony A. Atkinson, and Huntley W. H. Zia.
818. Some Basis Theorems for Integral Monoids. Robert G. Jeroslow.
819. Protective Planning-Horizon Procedures for a Deterministic Cash Balance Problem. J. Mensching, S. Garstka, and T. Morton.
820. Market Anticipations, Rational Expectations and Bayesian Analysis. Robert M. Townsend.
821. A Converse for Disjunctive Constraints. C. E. Blair and R. G. Jeroslow.
822. The Conduct of Monetary Policy Under Current Monetary Arrangements. Allan H. Meltzer.
823. Infinite-Horizon Dynamic Programming Models — A Planning-Horizon Formulation. Thomas E. Morton.
824. Algoristics for Single-Machine Sequencing with Precedence Constraints. Thomas E. Morton and Bala Ganga Dharan.
825. The Effects of Consumer Education on Consumer Product Safety Behavior. Richard Staelin.
826. The Roles for Research and Development in Auditing. Robert S. Kaplan.
827. Comments on "The Commission on Auditor's Responsibilities: Report of Tentative Conclusions." Robert S. Kaplan. Stanford Lectures in Accounting 1977.
828. Christmas Toy Manufacturer's Problem: An Application of the Stochastic Maximum Principle. Suresh P. Sethi and Gerald L. Thompson.
829. Hidden Z-Matrices with Positive Principal Minors. Jong-Shi Pang.
830. Some Theorems on the Core of a Non-Sidepayment Game with a Measure Space of Players. T. Ichiishi and Sh. Weber.
831. Disjunctive Programming. Egon Balas.
833. The Information Content of Financial Accounting Numbers: A Survey of Empirical Evidence. Robert S. Kaplan.
834. Consumer Product Safety: Strategies for Reducing the Incidence of Product-Related Injuries. Richard Staelin and R. D. Pittle.
835. Effects of Changes in Job Characteristics on Work Attitudes and Behaviors: A Naturally Occurring Quasi-Experiment. J. Richard Hackman, Jone L. Pearce, and Jane Caminis Wolfe.
836. Conditions Under Which Employees Respond Positively to Enriched Work. Greg R. Oldham and Jone L. Pearce.
837. The Price-Level Restatement and Its Dual Interpretation. Yuji Ijiri.

(continued on inside back cover)

ON PARSIMONIOUS EXPLANATIONS OF PRODUCTION RELATIONS

Herbert A. Simon

Carnegie-Mellon University, Pittsburgh, Pennsylvania, USA

Abstract

This paper examines three bodies of macroscopic evidence that are relevant to testing the classical theory of production. No one of these bodies of evidence offers any persuasive support for the classical theory. Fits to data of the Cobb-Douglas and ACMS functions appear to be artifactual, the data actually reflecting the accounting identity between values of inputs and outputs. Similarly, there is little or no support for the U-shaped long-run cost curve, the data on firm size distributions being better explained by the Gibrat hypothesis and constant costs. Finally, observed distributions of executive salaries are explainable in simple sociological terms, without recourse to classical theory.

Introduction

In testing theories aimed at explaining empirical phenomena, it is not enough to satisfy ourselves that the observed data are consistent with the theory. We must also ask whether the data can be explained equally well by other, perhaps weaker and simpler theories. In a recent paper (Simon, 1979), I expressed some reservations as to how far macroscopic empirical data support the classical theory of the firm in competition with other theories that might be put forward to explain these same data. In the present article, I should like to expand my rather brief remarks in the earlier paper and examine in some depth a small number of macroscopic phenomena that have been regarded as providing particularly significant support for the classical and neoclassical theory.¹

Economics has been prone to using highly aggregated data to test theories that are nevertheless founded upon postulates of the behavior of individual human beings or groups of human beings organized as business firms. This mode of inquiry may be highly efficient if it works—avoiding both the neces-

¹ This paper has been in gestation for a little more than a decade. During that time, in discussions and correspondence on these topics, I have incurred numerous intellectual debts to Martin Bronfenbrenner, Richard Cyert, Yuji Ijiri, Dale Jorgenson, Robert Lucas, Timothy McGuire, and Robert Solow. None of them, of course, are responsible for the positions I take here.

sity of tedious field studies, and the troublesome problems of aggregating data from samples of individual firms. If, after inquiry, we were to agree that the assumptions of the classical theory had been tested adequately at the macroscopic level, then we could continue to do business as we had been doing it in the past. Regression analysis of national data could continue to serve as the principal tool of empirical research in economics. However, if we find the existing empirical tests of the theory to be inadequate, we shall have either to invent novel tests at the macroscopic level or to turn to more microscopic investigations, and while doing so solve the methodological problems that they pose.

I. Production Functions and Labor's Share

One key piece of the macroscopic evidence consists of inferences drawn from fitting the Cobb-Douglas (C-D) function (or, more recently, the ACMS production function) to cross-section and time series data. Some years ago, Ferdinand Levy and I undertook to show that the fit of the C-D function was, under certain conditions at least, a statistical artifact (Simon & Levy, 1963). But neither this nor earlier criticisms of the interpretations of the C-D function (see especially, Mendershausen, 1938; Reder, 1943; Phelps-Brown, 1957) seems to remove those interpretations from the field. The fact that fitted C-D functions commonly "predict" correctly labor's share of product continues to be cited as a major piece of macroscopic evidence for the classical theory.

My own interest in this question was revived about a decade ago by claims of Jorgenson & Siebert (1968) that evidence cited by Walters (1963) in his survey of production and cost functions settled the issue once and for all in favor of the classical theory. Their evaluation of Walters' survey is worth quoting, as showing how conclusions become firmer as they filter from secondary to tertiary sources.

The empirical evidence is so largely favorable to (the classical) theory that current research is concentrated on such technical questions as the appropriate form for the production function and the statistical specification of econometric models of production. A recent survey of the literature on cost and production functions by Walters lists 345 references, almost all presenting results of econometric tests of the classical theory of the firm.

My curiosity aroused, I returned to the Walters paper for a closer look at the evidence. The paper, and the articles it cites, refer to two tests, and only two tests, of the classical theory. The first concerns the equality of the labor share of income with the labor exponent of fitted C-D functions—an equality implied, of course, by the theory. The second concerns the shape of short-run and long-run cost curves—whether they are U-shaped as classical competitive theory requires.

Most of the 345 papers cited in Walters' bibliography are not concerned with either of these tests, but simply with estimating just-identified production functions or discussing estimation methods. Hence, the Walters evidence seemed far less overwhelming than the number of his references had suggested. However, we cannot assess the adequacy of theories by counting footnotes. What I should like to do here is to consider the logical and theoretical status of the production function evidence.

I must first review briefly the reason why the data on production functions (Cobb-Douglas and ACMS) has been thought to provide significant support for the classical theory. The C-D function has now served for 45 years as a useful tool for representing empirical relations between capital and labor inputs, on the one hand, and outputs of product, on the other. The observation that we can fit C-D functions to data, using just-identified regression models, proves nothing, of course, about the validity of the classical, or any other, theory of the firm. It is of the nature of a just-identified model that it can always be fitted (more or less well) to the observed data, and that the fitted regressions will provide unique estimates of the model's parameters. A just-identified model, taken by itself, is a tool for estimating parameters of an accepted theory, not a tool for testing theories. Unless a model is over-identified, so that data could be inconsistent with it, it cannot contradict the theory, hence cannot test it.

The fact, therefore, that a considerable part of the variance of outputs can be explained by fitted C-D functions has no implications for the question before us. What *do* have implications are two further empirical generalizations that can be induced from the many such parameter-estimating studies to be found in the literature:

1. When the exponent of the labor factor of the fitted C-D function is compared with the labor share of income, there is usually reasonably good agreement between these two numbers; and
2. In a majority of cases, the sum of the exponents of the labor and capital factors of the fitted function is close to unity—the fitted C-D functions are very nearly homogeneous of the first degree.

The theoretical import of these findings was very early noted by econometricians, and excited considerable interest, along with understandable skepticism. Were these results indeed to be attributed to the operation of a classical competitive system, or was there a possibility that they might be artifactual?

Most of the early C-D studies used time series data. Mendershausen (1938) showed that the results were dubious by reason of the collinearity of the data. His argument was countered by the demonstration that essentially the same results (first degree homogeneity, and agreement of labor's exponent with labor's share) could be obtained with cross-section data.

A few years later, Reder (1943) raised an even more damaging objection, showing that the fitted Cobb-Douglas function could not be a firm production function in the ordinary sense, but only an envelope of such functions. What also followed from his analysis, but what he does not seem to have made explicit, is that if different firms on the same C-D production surface employ different ratios of capital to labor, they must have different wage rates. But in a competitive economy, all the firms or industries in a cross-section study must pay the same wage. Hence, the assumption that all the observations in such a study lie on the same production surface requires market imperfections of some kind.

Bronfenbrenner (1944), who had participated with Douglas in some of the early empirical work, replied to Reder in an interesting way. He agreed that the C-D function was probably not a production function in the classical sense, but an "interfirm function", and that it might be just as well to express the value of product as a *linear* function of labor and capital inputs. In what follows, I will have occasion to take up and develop further the ideas of both Mendershausen and Bronfenbrenner.

If we are not, in fact, observing the production function, how can we explain the observed equality, with fitted C-D functions, between the labor exponent and labor's share of income? A possible answer, consistent with Bronfenbrenner's interpretation, was proposed by Phelps-Brown (1957) and developed formally by Simon & Levy (1963). The latter authors, however, carried through the derivation only for the special case where the C-D function was already assumed to be homogeneous of the first degree, and their analysis holds only for cross-section data. These results are extended in this paper to the general case of the C-D function, to the ACMS production function, and to time-series data.

In recent years, most work on the C-D function has focused on statistical methods for estimating its parameters. Most empirical studies have used simple least-squares as the estimation method, even though extremely serious problems of identification are ignored by this method. In fact, research on the estimation problem, beginning with the classic paper of Marschak & Andrews (1944), has led to only gloomy conclusions about the possibilities of obtaining satisfactory estimates of the parameters from macroscopic data (see the review by Nerlove, 1965; and the discussion in Walters, 1963). A fair conclusion is that the parameters that have been estimated empirically in cross-section and time series data have little or nothing to do with the classical concept of a production function.

However, if we accept this pessimistic conclusion, we are again left with the mystery. Why does the estimated labor coefficient agree with independently measured data on labor's share of the product? And why do the estimated labor and capital coefficients usually add to unity, or something very close? In the next sections, I shall attempt to solve this mystery.

II. Cross-Section Data

The general Cobb-Douglas function can be written:

$$P = KL^g C^h, \quad (1)$$

where P is the *value* of output (usually, in fact, the value added), L is quantity of labor employed, C is value of capital employed, and K , g , and h are constants.

Differentiating, we obtain the marginal productivities of labor and capital, respectively:

$$\partial P / \partial L = gP / L \quad (2a)$$

$$\partial P / \partial C = hP / C. \quad (2b)$$

Since, in classical theory, these marginal productivities must equal wages, w , and the cost of capital services, r , respectively, it follows that:

$$w = gP / L \quad (3a)$$

$$r = hP / C. \quad (3b)$$

However, if labor's share of product, wL , plus capital's share, rC , equals total product, then:

$$wL + rC = gP + hP = P, \quad (4)$$

from which it follows immediately that

$$g + h = 1. \quad (5)$$

Thus, classical theory combined with the assumption that the production function is C-D demands (equation 3) the equality observed in the empirical data between g and labor's share of income, wL/P . Further, the C-D function must be homogeneous of the first degree.

III. Statistical Estimation Problems

At this point it will be helpful to insert some comments on the problem of estimating a C-D production function. I will not try to review the large and sophisticated econometric literature that has grown up around this topic (see Nerlove, 1965), but simply point out the central issue. I will limit myself, also, to the simplest case, that of interfirm cross-sectional data.

Consider a perfectly competitive economy. If a sample of firms were all producing on a function satisfying equation (1), and all paying the same wages and costs of capital, then, by equations (3), all of our observations of the value of their output and their labor and capital inputs would exhibit the same P/L and C/L ratios, and there would consequently be collinearity in the data. Of

course, if all the firms have the same long-run cost curve, matters are even worse, for then the scatter of values of P should be small.

In order, then, to secure identifiability, one or another kind of independent disturbance must be introduced into the equations. The disturbances may represent market imperfections leading to differences in prices of inputs or outputs, or differences in the parameters of the production functions resulting from varying "managerial quality", or other unobservables. As Marschak & Andrews (1944) demonstrated, the estimates of the parameters are highly sensitive to changes in these underlying assumptions.

For these reasons, it is hard to place much confidence in estimates of g and h arrived at through ordinary least-squares, although that is the source of most of the parameters in the literature. And that leads to the mystery we are trying to solve here.

IV. The Interfirm Function

If we are not dealing with a genuine production function—and it seems pretty clear that we are not—why are we able to fit the data? What is the artifact that makes this possible?

Suppose that the true relation among P , L and C is Bronfenbrenner's "interfirm function," an accounting equation that asserts that value added must equal the sum of the cost of labor and the cost of capital:

$$P = wL + rC. \quad (6)$$

Suppose, further, that a Cobb-Douglas function, (1), is fitted mistakenly to data that were actually generated by equation (6). What can we infer about the parameters of the mistakenly fitted Cobb-Douglas function? If we used least-squares or some similar method, the fit would depend on the distribution of observations in the plane of (6). Let us apply a different method that does not depend on that distribution.

We consider the function (1) that is tangent to (6) somewhere near the center of all the observations. This surface will not necessarily provide the best fit, in a least-squares sense, of (1) to the observations on (6) (and we will consider presently how to improve the fit), but it will be an approximation. Let us designate the coordinates of the point of tangency by P_0 , L_0 , C_0 . Then, from (6),

$$\partial P / \partial L = w; \quad \partial P / \partial C = r. \quad (7)$$

Combining (2) with (7), under the above assumption of tangency, we find:

$$gP_0/L_0 = w; \quad hP_0/C_0 = r, \quad (8)$$

hence,

$$(gP_0/L_0)L_0 + (hP_0/C_0)C_0 = P_0. \quad (9)$$

and consequently,

$$g + h = 1, \quad (10)$$

so that the fitted equation (1) is homogeneous of first degree.

Finally, solving (8) for g , we obtain:

$$g = wL_0/P_0, \quad (11)$$

so that the labor coefficient of the fitted Cobb-Douglas function is equal to labor's share of income. But now these results follow from the simple accounting identity, (6), without any necessity for the classical assumption, (3), that the wage rate equals the marginal productivity of labor.

The alternative explanation gains added credibility from the implication of (10) that the erroneously fitted C-D function will be homogeneous of the first degree—that it will exhibit constant returns to scale for capital and labor jointly. The empirically observed fact that, with cross-section data for a single country, the sum of the two exponents is usually close to unity—sometimes a little smaller, sometimes a little larger—has always been something of an embarrassment for the classical theory. For, as is well known, decreasing returns to scale are essential to competitive equilibrium. Our alternative explanation produces the observed result without implying anything about whether returns to scale are decreasing, constant, or increasing. Hence, it is compatible with competitive equilibrium, while the marginalist explanation is not whenever the sum of the coefficients equals or exceeds unity.¹

Our result depends, however, on an unusual curve-fitting method—taking the Cobb-Douglas function that is tangent to the plane of the observations. It may readily be seen that if the Cobb-Douglas function is homogeneous of the first degree, it will lie entirely below the plane of observations except along the ray from the origin to the point of tangency. Still, the approximation, evaluated in Column (2) of Table 1, is quite close. This table shows the ratio of estimated (P_e) to actual (P) output for different values of L/C . The estimated values generally lie within 10 or 15 per cent of the actual values.

The fit of function to data can be improved by increasing the constant, K , in equation (1), so that it will overestimate some observations while underestimating others. In general, the function will then overestimate P when L/C is close to L_0/C_0 , and underestimate P when L/C is far from the mean value.

The relative error of estimate $P_E/P = (P_{est}/P_{act})$ is given by the ratio of the right-hand sides of (1) and (6), respectively:

$$P_E/P = (KL^gC^h)/(wL + rC) = K(L/C)^g/(w(L/C) + r). \quad (12)$$

¹ Jorgenson (1974), in reviewing a wide range of empirical data on production functions concludes that the evidence consistently supports constant returns to scale. He describes there the kind of dynamic investment function that is consistent with this condition. Of course, constant returns pose no theoretical problems under conditions of imperfect competition where inputs and outputs have finite price elasticities.

Table 1. *Ratios of estimated P_E (Equation 1) to observed P (Equation 6) for varying ratios of L/C* P_E —Eq. (6) tangent to Eq. (1). P'_E = Col. 2 $\times$ 1.11. P''_E = Col. 2 $\times$ 1.08

(1) L/C	(2) P_E/P	(3) P'_E/P	(4) P''_E/P
0.20	0.77	0.86	
0.25	0.83	0.92	0.90
0.33	0.89	0.99	0.96
0.50	0.96	1.07	1.04
1.00	1.00	1.11	1.08
2.00	0.97	1.08	1.05
3.00	0.93	1.03	1.00
4.00	0.89	0.99	0.96
5.00	0.86	0.96	

Hence, P_E/P is a function of L/C , and the goodness of approximation will depend only on that ratio. Choosing our units so that $L_0/C_0=1$, Column 2 of Table 1 shows the value of E for a range of values of L/C from 1/5 to 5/1. In Column 3 of Table 1, adjusted values of P_E/P are shown where the value of K has been multiplied by a factor of 1.11 to provide a better fit.

In the table, we have allowed an extreme range of 25 to 1 from the lowest to the highest ratio of L to C (from 0.2 to 5). Even over this wide range, far wider than any encountered in the literature, the fitted C-D function approximates the data on the plane (6) with a maximum error of less than 15 per cent. Column 4 of Table 1 indicates the relative error when the C-D function is fitted to the plane over a range of L/C of 16 to 1. Now the greatest error is only 10 per cent. Since in the data actually observed, most of the sample points lie relatively close to the mean value of L/C , we can expect *average* estimating errors of less than 5 per cent.

V. The ACMS Production Function

In the last one and one-half decades, the Cobb-Douglas function has been to some extent supplanted in econometric work on production by the more general ACMS production function. The ACMS function is given by:

$$P = \gamma(\delta C^{-\epsilon} + (1-\delta)L^{-\epsilon})^{-(1/\epsilon)}. \quad (13)$$

If we assume, again, that the interfirm function (6), is the true function relating P , L , and C , but that it has been replaced mistakenly by a function of the form (13), tangent to it at the mean values, P_0 , L_0 , C_0 , then, differentiating (13) and using (7), we obtain:

$$\begin{aligned} (\partial P/\partial L)_0 &= (1-\delta)(P_0/L_0)^{(1+\epsilon)} = w; \\ (\partial P/\partial C)_0 &= \delta(P_0/C_0)^{(1+\epsilon)} = r. \end{aligned} \quad (14)$$

Combining (14) with (6), we find:

$$(1-\delta)(P_0/L_0)^{(1+\varrho)}L_0 + \delta(P_0/C_0)^{(1+\varrho)}C_0 = P_0, \quad (15)$$

whence,

$$(1-\delta)(P_0/L_0)\varrho + \delta(P_0/C_0)\varrho = 1. \quad (16)$$

Equation (16) will be satisfied identically if $\varrho=0$. But it can be shown (Arrow et al., 1961, p. 231) that as ϱ goes to zero in the limit, (12) becomes simply:

$$P = \gamma L^{(1-\delta)} C^\delta \quad (17)$$

i.e., it becomes the special case of the Cobb-Douglas function, (1), for which $g+h=1$.

Hence, if we were to fit, mistakenly, the ACMS function, (13), when the true relation among the variables was (6), we should find that $(1-\delta)$ and δ equalled the labor and capital shares, respectively, and that ϱ was zero, or close to it.

The ACMS function has been fitted, and $\sigma=1/(1+\varrho)$ estimated for a wide variety of data. These studies have been reviewed by Jorgenson (1974), who finds that, when good statistical methodology is used, the estimated values of σ are quite close to unity, implying that ϱ is close to zero, and hence that the conditions for a C-D function are satisfied, and consequently also the conditions for the accounting relation (6).

Thus, for the ACMS function as for the Cobb-Douglas function, the fact that the observed labor and capital shares are close to the values required by the theory is readily explainable as an artifact, without the need for marginalist assumptions; the alternative assumption is simply that the data being fitted really represent the relation (6). Moreover, the observed values of ϱ , not predicted by the marginalist theory, are close to those we would obtain with the alternative model. We cannot, therefore, regard any of these empirical studies—whether of the Cobb-Douglas or the ACMS functions—as providing aid and comfort, much less validation, for the neo-classical theory of the firm.

VI. Time-Series Studies

The argument of the previous pages does not carry over satisfactorily to the case of time-series data; for the interfirm function assumes that wages are constant over all observations, an assumption unlikely to be valid for time series of any great length. But in these cases, the surprisingly good fit of the Cobb-Douglas function to the data can also be explained as artifactual by an extension of the argument of Mendershausen (1938).

We need to introduce one additional assumption, which may be acceptable simply because it fits the facts well, but which may also be given a theoretical interpretation. It is generally the case that the ratio of P to C is almost constant over time (see, e.g., Solow, 1970, pp. 2–3, 5–8):

$$C = sP, \quad (18)$$

where s is a constant.

Theories of saving like the Modigliani–Brumberg life-cycle savings hypothesis (1954) predict something like this constancy. In such theories, saving is a means of smoothing consumption over the life cycle. For a given age distribution of the population, fixed retirement age, stable age-earnings profile, and constant interest rate, this smoothing requires, in the aggregate, a fixed ratio of savings to total output.¹ But whether derived from theoretical considerations or not, equation (18) holds remarkably well for a wide variety of empirical data, including the time series data that have been used to fit the C-D function.

Suppose, now, that we have the accounting identity:

$$P(t) = w(t)L(t) + rC(t) = w(t)L(t) + rsP(t), \quad (19)$$

where the wage, w , is no longer a constant but a function of time. In particular, suppose that $w(t)$ and $L(t)$ are exponential functions

$$w(t) = w_0 e^{at}; \quad (20a)$$

$$L(t) = L_0 e^{bt}. \quad (20b)$$

Substituting (20) in (19), and writing P_0 for $P(0)$, we obtain:

$$P(t) = (P_0/L_0) e^{at} L(t) = P_0 e^{(a+b)t}. \quad (21)$$

Hence, under these assumptions, P will grow exponentially, at a rate that is the sum of the rates of growth of wages and the labor force.

Now in fitting the C-D function to time-series data, an exponential growth factor is usually introduced to allow for technological change. Assume that we fit a C-D function, so modified, to data generated by relations (19) and (20). The function replacing (1), will be of the form:

$$P_* = K e^{ct} L^g C^h, \quad (22)$$

where c is the parameter describing technological change. If (18) holds, so that $C = sP$, then by substituting this relation in (22), we readily derive:

$$P_* = B e^{ct} L^g P^h, \quad \text{where } B = K(rs)^h \quad (23)$$

$$P_*^{(1-h)} = B e^{ct} L^g, \quad \text{whence} \quad (24)$$

$$P_* = A e^{ct/(1-h)} L^{g/(1-h)}, \quad A \text{ a constant.} \quad (25)$$

¹ Of course the constants listed here need only be grossly stable to satisfy our curve-fitting requirements. Solow (1970, pp. 3–8) concludes that the empirical data in fact exhibit such stability, at least to a first approximation.

Comparing this with (21), we see that (25) will be satisfied identically in t if $A = P_0/L_0$, $c/(1-h) = a$, and $g/(1-h) = 1$, the latter condition implying that $g+h=1$. Hence, c , the estimated rate of technological advance, will equal $(1-h)$ times the rate at which the average wage increases; while the fitted C-D function will be homogeneous of the first degree.

The results of this section also explain the good fit obtained by Arrow, et al. (1961) of the ACMS function to international industry data. Here, as with time series, we cannot assume w to be constant across observations. Essentially, Arrow et al., fit the relation:

$$\log (P/L) = a \log (w), \quad (26)$$

obtaining good fits with a generally a little less than 1.0. But this is implied by the accounting identity (19), which follows on the assumption (18), as before. The relation (18) holds empirically to an excellent approximation for international comparisons.

VII. Summary: The Cobb-Douglas Function

Empirical data on the Cobb-Douglas and ACMS production functions have been alleged to provide substantial support for the classical theory of the firm—so substantial that further testing of that theory, as distinguished from elaboration of its detail, was no longer necessary. An examination of the evidence suggests instead that the observed good fit of these functions to data, the near equality of the labor exponent with the labor share of value added, and the first degree homogeneity of the function are very likely all statistical artifacts. The data say no more than that the value of product is approximately equal to the wage bill plus the cost of capital services. This interpretation of the statistical findings is plausible for both interindustry cross-sectional studies and time-series studies, the latter for either a single industry or a whole economy.

VIII. On Constant Returns to Scale

The second issue that Walters examined in his survey of empirical tests of the theory of production was what the data revealed about the shape of the long-run cost curve. If the studies of production functions he surveyed fail to validate classical theory with perfect competition, the empirical data on cost curves that he cited are no more helpful to that end. To indicate why this is so, we cannot do better than quote Walters himself (1963, pp. 40–41):

To summarize, the theoretical arguments suggest that the short run average cost curve has the typical U shape, although Menger has added several reservations. Theory is reasonably clear on the proposition that long run average costs

may be expected to decline at first with increasing scale. But for high outputs the theoretical arguments do not seem to be so convincing. Choice between the alternatives must depend on the empirical evidence.

Of course, if empirical evidence must decide the actual shape of the long run cost curve, and if any shape that evidence may disclose is consistent with neoclassical theory, then that same evidence is useless to validate or invalidate the theory. But the situation is not much better with respect to the short run curve. While (subject to the reservations mentioned) empirical evidence that the actual curves are *not* U-shaped might be taken as reason for distrusting the theory, logic does not force the converse on us. If the actual curves *are* U-shaped, this finding is consistent with the classical theory, but does not imply its validity.

We need hardly debate, however, what such a finding would imply, since the empirical data, as summarized by Walters, are so inconclusive. After reviewing the evidence, Walters is able to make no stronger defences of the U-shaped cost curve than these:

... the evidence in favour of constant marginal costs is not overwhelming. Certainly the revision of theory to include this phenomenon is not an urgent matter (Walters, 1963, p. 51).

... for "competitive" industries, the U-shaped hypothesis does not inspire great confidence. But this is not because it has been refuted by direct empirical evidence. On the contrary, this lack of faith is occasioned by the very few *opportunities* for collecting evidence to refute the theory directly. Instead we are driven to indirect and circumstantial evidence. But at least there is no large body of data which convincingly contradicts the hypothesis of a U-shaped long run cost curve and the fruitful results which depend on it (Walters, 1963, p. 52).

Thus, Walters' defense of classical theory, resting on evidence from fitted C-D production functions and estimated cost curves, ends, not with a ringing conclusion that the theory has been vindicated, but with the mild claim that the data on cost curves, though generally not supportive of the U-shaped curve usually thought to be necessary for classical competitive theory, are not so damning as to require the theory to be abandoned!

Of course, if we give up the assumption of perfect competition, then constant or decreasing costs create no problem for theory. Jorgenson himself, in a new review of the evidence (Jorgenson, 1974), concludes that the data in most cases point strongly to constant costs, and in this and other papers has developed an investment theory compatible with that finding.

IX. Constant Costs and Firm Sizes

A form of data that provides indirect evidence about the shapes of cost curves, but which was not examined in Walters' article, are the statistics of business firm size distributions. The classical theory, which derives the U-shaped long

run cost curve as an envelope of U-shaped short run cost curves (Ijiri & Simon, 1977, pp. 7-11), proves either too much or too little. For if it assumes that all firms in an industry are faced with the same cost structure, then it predicts, grossly contrary to fact, that all firms in the industry will be of the same size. But if it assumes different costs for different firms, then it makes no prediction at all about the size distribution. The empirically observed size distributions, on the other hand, exhibit a remarkable regularity—they are all highly skewed and can be fitted closely by the Pareto distribution or very similar skew distributions. The evidence is reviewed in Ijiri & Simon (1977).

Firm size distributions resembling those actually observed are readily generated by any one of a number of stochastic models, all of which have in common some form of the Gibrat hypothesis. By this hypothesis, probabilities of growth are independent of current size—a postulate that can, in turn, be derived from the assumption of constant returns to scale. The latter assumption is consistent not only with the observed firm size distributions, but also with the fact (Singh & Whittington, 1968; Steindl, 1965) that there is little or no correlation between firm sizes and profit rates.

Thus, classical theory is either incompatible with, or irrelevant to, the observed distributions of business firm sizes; while a simple stochastic growth theory, making no assumptions of profit maximization, does a good job of predicting the actual size distributions.

Some attempts have been made to account for the observed skew distributions in terms of classical theory, but those with which I am familiar either fall short of the mark or require *ad hoc* assumptions that are not especially plausible. I will mention one recent attempt falling in each of these two categories.

Hjalmarsson (1974) has developed an interesting theory to account for the size distributions of establishments and firms on the assumptions of a market that grows steadily at a geometric rate, and of increasing returns to scale. The periodic investment in new plant will then produce a skew size distribution of plants similar to the observed distributions. This model gives a not implausible account of one of the factors that may account for the wide spread in *plant* sizes in most industries, which is all Hjalmarsson claims for it. It is hard to see how it could be extended to a model of firm sizes, except in those rare cases where each firm does its manufacturing in a single establishment.

Robert E. Lucas (1978) develops an entirely different explanation of the distribution of firm sizes, based on the assumption that there is a distribution in human ability to manage assets effectively, and consequently, a distribution in the assets that are entrusted (by market mechanisms) to each manager. Unfortunately, the distribution of firm sizes that will result from the operation of this process depends sensitively on the distribution of managerial ability in the population. Hence, Lucas' model explains the observed phenomena—

the firm size distributions—by postulating unobserved phenomena—the distribution of managerial abilities, which are fully as complex as those explained.

There are additional serious difficulties with Lucas' theory, one of which he mentions. The theory predicts that top executive salaries will be proportional to firm sizes, whereas the data show clearly that they vary only as a fractional power (approximately cube root) of firm sizes. But the latter relation is also readily predictable from a non-classical non-maximizing model that assumes (1) that business firms are regular hierarchies, with a narrow range in the number of employees per immediate supervisor, and (2) that there is a socially accepted "normal" ratio of the salaries of supervisors to the salaries of their immediate subordinates (Simon, 1957).

For let C be the total annual compensation of the highest paid official in a firm with annual dollar sales, S , and a and k constants. The observed empirical relation is: $C = kS^a$. Suppose that each executive has exactly n immediate subordinates, and that his salary is b times the salary of each such subordinate. Then, if L is the number of levels in the organization, the total sales, assumed proportional to total employment, will be about $S = n^L$. Likewise, the highest paid official will receive $C = Bb^L$. Taking logarithms, eliminating L between the two equations, and solving for C in terms of S , we get $\log C = a \log S + \log B$, where $a = (\log b)/(\log n)$.

Lucas' theory also fails to explain under what circumstances a set of people will be organized in an independent hierarchy (a firm), and under what circumstances they will simply form a subhierarchy directed by a subsidiary manager. On all accounts, Lucas' theory is neither parsimonious nor effective in predicting the observed facts. But we have just seen that these same facts—the distributions of both firm sizes and salaries—are readily predicted on the base of relatively weak nonclassical stochastic assumptions.

X. Summary

In this paper, I have referred to three bodies of macroscopic evidence that are relevant to testing the classical theory of the firm, and which are based on econometric techniques of the sorts that economists are most familiar and comfortable with. No one of these three bodies of evidence offers any persuasive support to the classical theory.

Most of the data on the relation between outputs and labor and capital inputs ("production functions") is not inconsistent with the assumption of constant returns to scale, or with the data showing that labor's share in the value of output is predicted correctly by the parameters of the fitted Cobb-Douglas and ACMS functions. On the other hand, grave questions of econometric method have been raised (and not answered) about the legitimacy of regarding the fitted functions as genuine production functions, in the meaning that classical theory attaches to that term. Since the observed phenomena

can as readily be explained on the weaker assumption that what is being observed is simply the accounting relation equating value of output to the sum of factor costs, the criterion of parsimony would lead us to prefer the latter explanation to the classical one.

Similarly, the empirical study of cost functions has provided little or no support for the classical theory of a U-shaped long run cost curve that is the envelope of U-shaped short run curves. Here the data are better explained by a simple hypothesis of constant costs that also leads, via the Gibrat principle, to an effective theoretical explanation of the observed distributions of business firm sizes.

In this paper, I have not, of course, reviewed all the kinds of macroscopic evidence that might be applied to testing the validity of the classical assumptions. In my previous paper (Simon, 1979), I also discussed briefly the evidence from the fact that fitted demand curves are usually negatively sloping. Going farther afield, one might consider to what extent the types of neoclassical investment functions that have been fitted to firm accounting data provide support for classical theory. There are probably other forms of evidence in addition to these. But I hope that the present paper, while not covering the entire ground, has reviewed enough of the evidence usually advanced in support of the classical theory to show that the support usually assumed is not really there.

If the classical theory, even though not supported by evidence, were the only one known to us, we might wish to hold to it until a better came along. However, in the cases I have examined here, simpler, more parsimonious theories are available, which explain the data as adequately or more adequately, and which avoid the counterfactual assumptions of global human rationality that make the classical theory so implausible.

References

- Arrow, K. J., Chenery, H. B., Minhas, B. S. & Solow, R. M.: Capital-labor substitution and economic efficiency. *The Review of Economics and Statistics* 43, 225-50, 1961.
- Bronfenbrenner, M.: Production functions: Cobb-Douglas, interfirm, intrafirm. *Econometrica* 12, 35-44, 1944.
- Hjalmarsson, L.: The size distribution of establishments and firms derived from an optimal process of capacity expansion. *European Economic Review* 5, 123-40, 1974.
- Ijiri, Y. & Simon, H. A.: *Skew distributions and the sizes of business firms*. North-Holland, Amsterdam, 1977.
- Jorgenson, D. W.: Investment and production: a review. In M. D. Intriligator and D. A. Kendrick (eds.), *Frontiers of quantitative economics*, vol. II, pp. 341-75. North-Holland, Amsterdam, 1974.
- Jorgenson, D. W. & Siebert, C. D.: Optimal capital accumulation and corporate investment behavior. *Journal of Political Economy* 76, 1123-51, 1968.
- Lucas, R. E., Jr: On the size distribution of business firms. *The Bell Journal of Economics* 9, 508-23, 1978.

- Marschak, J. & Andrews, W. H. Jr: Random simultaneous equations and the theory of production. *Econometrica* 12, 143-205, 1944.
- Mendershausen, H.: On the significance of Professor Douglas' production function. *Econometrica* 6, 143-53, 1938.
- Modigliani, F. & Brumberg, R.: Utility analysis and the consumption function: an interpretation of cross-section data. In K. K. Kurihara (ed.), *Post Keynesian economics*, pp. 388-436. Rutgers Univ. Press, New Brunswick, N.J., 1954.
- Nerlove, M.: *Estimation and identification of Cobb-Douglas production functions*. North-Holland, Amsterdam, 1965.
- Phelps-Brown, E. H.: The meaning of the fitted Cobb-Douglas function. *The Quarterly Journal of Economics* 71, 546-560, 1957.
- Reder, M. W.: An alternative interpretation of the Cobb-Douglas function. *Econometrica* 11, 259-64, 1943.
- Simon, H. A.: The compensation of executives. *Sociometry* 20, 32-35, 1957.
- Simon, H. A.: Rational decision making in business organizations. *American Economic Review*, 69, 493-513, 1979.
- Simon, H. A. & Levy, F. K.: A note on the Cobb-Douglas function. *Review of Economic Studies* 30, 93-4, 1963.
- Singh, A. & Whittington, G.: *Growth, Profitability and valuation*. Cambridge Univ. Press, Cambridge, 1968.
- Solow, R. M.: *Growth theory*. Oxford University Press, Oxford, 1970.
- Steindl, J.: *Random process and the growth of firms: a study of the Pareto law*. Hafner, New York, 1965.
- Walters, A. A.: Production and cost functions: an econometric survey. *Econometrica* 31, 1-66, 1963.

838. From Accounting to Accountability: Steps to a Corporate Social Reporting. W. W. Cooper and Yuji Ijiri.
839. Financial Distress in Private Colleges. Katherine Schipper.
840. Cash-Flow Accounting and Its Structure. Yuji Ijiri.
841. The Nonstationary Infinite Horizon Inventory Problem. Thomas E. Morton.
842. An Empirical Analysis of the Monetary Approach to the Determination of the Exchange Rate. Robert J. Hodrick.
843. Intermediation with Costly Bilateral Exchange. Robert M. Townsend.
844. Cutting-Plane Theory: Algebraic Methods. R. Jeroslow.
845. Sequential Investment Decisions with Bayesian Learning. R. M. Cyert, M. H. DeGroot, and C. A. Holt.
846. The Management of Universities of Constant or Decreasing Size. Richard M. Cyert.
847. An Arbitrage Model of the Term Structure of Interest Rates. Scott F. Richard.
849. Toward an Optimal Design of a Network Database from Relational Descriptions. Prabhuddha De, William D. Haseman, and Charles H. Kriebel.
850. Effects of Group Size, Problem Difficulty, and Sex on Group Performance and Member Reactions. Robert M. Bray, Norbert L. Kerr, and Robert S. Atkin.
851. On a Class of Least-Element Complementarity Problems. Jong-Shi Pang.
852. Bivalent Programming by Implicit Enumeration. Egon Balas.
853. Temporary Taxes as Macro-Economic Stabilizers. Walter Dolde.
854. A New Characterization of Real H -Matrices with Positive Diagonals. Jong-Shi Pang.
855. What Computers Mean for Man and Society. Herbert A. Simon.
856. How American Executives Disagree About the Risks of Investing in Eastern Europe. Ion Amariuta, David P. Rutenber, and Richard Staelin.
857. A Generalized Capital Asset Pricing Model. Scott F. Richard.
858. On the Solution of Some (Parametric) Linear Complementarity Problems with Applications to Portfolio Selection, Structural Engineering and Actuarial Graduation. J. S. Pang, I. Kaneko, and W. P. Hallman.
859. Leniency, Learning, and Evaluations. John Palmer, Geoffrey Carliner, and Thomas Romer.
860. Facets of One-Dimensional and Multi-Dimensional Knapsack Polytopes. Egon Balas.
861. Report of the Session on Branch and Bound/Implicit Enumeration. Egon Balas and M. Guignard.
863. Price Change Expectations and the Phillips Curve. Timothy W. McGuire.
864. On Estimating the Effects of Controls. Timothy W. McGuire.
866. A Bilinear-Quadratic Differential Game in Advertising. K. Deal, S. P. Sethi, and G. L. Thompson.
867. The Relationship Between Relative Prices and the General Price Level: A Suggested Interpretation. Alex Cukierman.
868. Judgment Based Marketing Decision Models: An Experimental Investigation of the Decision Calculus Approach. Dipankar Chakravarti, Andrew Mitchell, and Richard Staelin.
869. Heterogeneous Inflationary Expectations, Fisher's Theory of Interest and the Allocative Efficiency of the Bond Market. Alex Cukierman.
870. Product Reliability and Market Structure. Dennis Epple and Artur Raviv.
872. The Redistributive Effects of Inflation and of the Introduction of a Real Tax System in the U. S. Bond Market. Assa Birati and Alex Cukierman.
873. A Comparison of Rates of Return to Social Security Retirees Under Wage and Price Indexing. Robert S. Kaplan.
874. Epidemiology, Causality, and Public Policy. Lester B. Lave and Eugene P. Seskin.
875. A Dynamic Dominant Firm Model of Industry Structure. Finn Kydland.
876. Expectations and Money in a Dynamic Exchange Model. Milton Harris.
877. Bayesian Estimation and Optimal Designs in Partially Accelerated Life Testing. Morris DeGroot and Prem K. Goel.
878. On the Monetary Analysis of Exchange Rates. A Comment. Robert J. Hodrick.
879. The Elusive Median Voter. Thomas Romer and Howard Rosenthal.
884. Rational Expectations and the Role of Monetary Policy. A Generalization. Alex Cukierman.
885. Optimal Incentive Contracts with Imperfect Information. Milton Harris and Artur Raviv.
886. On the Use of Facet Analysis in Organizational Behavior Research: Some Conceptual Considerations and an Example. Zur Shapira and Eli Zevulun.

(continued on back cover)

The present series contains articles written by the faculty of the Graduate School of Industrial Administration. Publications began in the 1962-63 academic year and continue through to date. You may request copies and receive them from: Reprint Secretary, GSIA, Carnegie-Mellon University, Pittsburgh, Penna. 15213.

- 887. Writer-Based Prose: A Cognitive Basis for Problems in Writing. Linda Flower.
- 888. Issues and Models in Empirical Research on Aggregate Consumer Expenditure. Walter Dolde.
- 889. Rational Decision Making in Business Organizations. Herbert A. Simon.
- 890. Optimal Contracts and Competitive Markets with Costly State Verification. Robert M. Townsend.
- 891. Towards a Better Microeconomic Theory. Richard M. Cyert and Garrel Pottinger.
- 892. The Effects of Rate-of-Return Regulation on the Intensity of Use and Durability of Capital. Dennis Epple and Allan Zelenitz.
- 893. Comparison of Experiments and Information Measures. Prem K. Goel and Morris H. DeGroot.
- 894. Capital Allocation Within a Firm. R. M. Cyert, M. H. DeGroot, and C. A. Holt.
- 895. The Information Content of Discounts and Premiums on Closed-End Fund Shares. Rex Thompson.
- 896. Developing a Financial Planning Model for an Analytical Review: A Feasibility Study. Robert S. Kaplan.
- 897. Evaluating the Quality of Information Systems. Charles H. Kriebel.
- 898. Strengthening Cuts for Mixed Integer Programs. Egon Balas and Robert G. Jeroslow.
- 899. Dynamic Optimal Taxation, Rational Expectations and Optimal Control. Finn E. Kydland and Edward C. Prescott.
- 901. Cutting Planes from Conditional Bounds: A New Approach to Set Covering. Egon Balas.
- 902. Set Covering Algorithms Using Cutting Planes, Heuristics, and Subgradient Optimization: A Computational Study. Egon Balas and Andrew Ho.
- 903. Stein's Paradox and Audit Sampling. Yuji Ijiri and Robert A. Leitch.
- 904. Political Resource Allocation, Controlled Agendas, and the Status Quo. Thomas Romer and Howard Rosenthal.
- 909. Discussion. Scott F. Richard.
- 910. Comments on Externalities and Financial Reporting. Stanley Baiman.
- 911. The Cognition of Discovery: Defining a Rhetorical Problem. Linda Flower and John R. Hayes.
- 912. Four Essays on Procedural Rationality in Economics. Herbert A. Simon.
- 913. Pivot and Complement—A Heuristic for 0-1 Programming. Egon Balas and Clarence H. Martin.
- 914. Organization Capital. Edward C. Prescott and Michael Visscher.
- 916. Turnpike Horizons for Production Planning. Gerald L. Thompson and Suresh P. Sethi.
- 917. An Economics Approach to Modeling the Productivity of Computer Systems. Charles H. Kriebel and Artur Raviv.
- 918. Adjacent Vertices of the All 0-1 Programming Polytope. Egon Balas and Manfred W. Padberg.
- 920. Estimation of the Correlation Coefficient from a Broken Random Sample. Morris H. DeGroot and Prem K. Goel.
- 925. Should Accounting Standards be Set in the Public or Private Sector? Robert S. Kaplan.
- 926. A Competitive Theory of Fluctuations and the Feasibility and Desirability of Stabilization Policy. Finn Kydland and Edward C. Prescott.
- 927. Models of Money with Spatially Separated Agents. Robert M. Townsend.
- 928. The Limitations of Log-Linear Analysis. Howard Rosenthal.
- 929. Dynamic Effects of Government Policies in an Open Economy. Robert J. Hodrick.
- 932. A Dimension of the Myths and Science of Accounting. Robert S. Kaplan.



Management Information Systems: Progress and Perspectives (Carnegie Press, 1971).

Editors: C. H. Kriebel, R. L. Van Horn, and J. T. Heames. Price: \$11.50.

The Employment Game . . . Where Do You Fit? C. Douglas Mintmier, GSIA. Price: \$5.95