What Is Output?

PROBLEMS OF CONCEPT AND MEASUREMENT

ARTHUR B. TREADWAY

NORTHWESTERN UNIVERSITY

I. INTRODUCTION

IT SEEMS to me that the role of a theorist at a conference such as the present one is very difficult. He must be critical of existing empirical conventions and procedures or he surely has not earned his keep. On the other hand, he must be sufficiently constructive to suggest alternatives or he is likely to appear antiempirical or uninterested in reality. If he exercises his strongest natural propensity, to soar off into a visionary state of abstraction complete with page on page decorated with Greek notation, he is likely to find himself without an audience, and properly so. I take it then that he should start with first principles and be chary of remaining within the framework of contemporary empirical work while addressing himself to issues that could conceivably be dealt with empirically. As an avowed novice in the subject matter of this conference, I could hardly be constrained to any great degree by existing approaches, since only a mental Apicius could consume and digest in the short time allowed for this project all of the immense literature touching on the problems we are gathering to consider. I shall attend then to what I regard to be *first* principles: this seems to lead to at least a few empirically relevant statements. I shall, however, give only the most casually empirical basis for anything in the sequel.

What is output? Or, what is productivity? And why do we want to measure the latter? Aside from the few extreme advocates of pure inductive science among us, who might want to look at differential rates of growth of anything versus anything else just because these happen to be currently observable numbers, we can probably all agree that measuring output and productivity is part of our attempt to understand one of the foundation blocks of modern economic science, that is, technical conditions. To the extent that economics is an analytical science, it rests on the essential notions of: technical conditions, psychological conditions, the hypothesis of optimizing behavior, and hypotheses regarding social organization. We should want, therefore, to measure anything if and only if it sheds light on one or another of these categories. Now, we may want to measure output as an index of welfare; such a notion lies at the foundations of national income accounting.[1] Or we may want to measure output to be compared with other variables in an attempt to study technical relations.

It is a remarkable fact of the development of economics that we have recently exerted so much empirical effort in the study of apparently technical and, in demand studies, preference conditions. One might, on the basis of most of pre-Depression economics, have expected that the empirical study of technical and psychological relationships would have been relegated to the engineers and psychologists. It might appear that this would have left us perched precariously in the interstices of other people's pigeonholes. This view, however, neglects the nature of economics as a *social* science, a neglect no doubt too often validated by economists themselves. We are only too prone in both our theoretical efforts and in our empirical work to ignore the very *interactions between economic agents* that supply the prime raison d'etre for our separate intellectual discipline.[2] One of the main attractions of the theory of competitive equilibrium must be its implicit simplification of social interactions to the bare disembodied minimum; our work as economists would be vastly simpler if consumers did not depend on the good will of the baker or the butcher or the auto mechanic. Nowadays the efforts of some theorists are even going into finding social rules which, if adhered to, would lead one

[1] See Kuznets [21], chapter 1. Bracketed numbers cite references listed at the end of this paper.

[2] There are, no doubt, those who would be happy to define economics as the study of constrained optimization problems in deference to the one category above not already "allocated" to other disciplines. This remaining category would, however, go to the mathematicians if it were not for the social aspect of such problems in economics.

generation to be independent of the good will—or lack of it—of another! As for empirical research, one cannot help but be struck by the paucity of work done on issues involving nonmarket relationships between economic units. At the least, we know very little about such relationships.[3] At the worst, we ignore them as in the case of most productivity studies where equilibrium, competition, neutral tax structure and a host of other assumptions are made that sum up to one master hypothesis: the absence of relevant nonmarket interactions between economic agents.

I take it then that, if we are concerned with "output and productivity," we are measuring output in an effort to understand "technical" relationships.[4] Furthermore, I regard it as obvious that our prime concern is with the character of such relationships as embodying interactions between economic units whether or not these interactions take place in a market context; their nature is technical only in a broad sense.

Productivity measurement has been a popular activity of economists for a long time. Much attention has been devoted to rather substantial aggregates: the private sector, the agricultural sector, manufacturing, etc. Though productivity studies of some smaller aggregates are fairly old, we have recently reached a point where attention is being focused on economic units at many different levels of aggregation with attendant efforts to compare the patterns observed in these units. This conference is evidence of this pattern, focused as it is on the tertiary sector, a large and ill-defined aggregate, a sector that we realize to be particularly difficult to understand but that we somehow suspect to be highly relevant to future developments. Certain apparent empirical regularities stand out in the existing data: a massive shift in the pattern of employment from "goods" to "services" coupled with a substantial differential in measured productivity trends that favors "goods" and suggests that "services" may be somehow less productive.[5]

[3] Compare, for example, the relative bases in theory and fact available to Denison for estimating the effects on productivity change of "Economies of Scale and Increased Specialization," [6] pp. 173–181, versus changes in any one of the factor inputs measured by reference to market data and the assumption of competition.

[4] This is not to suggest that our efforts in this regard do not have implications for welfare, but that we conceptualize these relationships to be largely objective.

[5] These patterns were established by Fuchs [10], [11], [12] and, for countries in the OECD, by Lengellé [22]. A large part of service employment is in government,

One might add that we have also observed a significant increase in the ratio of so-called "nonproduction" workers to total employment in manufacturing, suggesting that some sort of employment shift to service type activities within the "goods" sector is taking place.[6] It must be apparent that such intersectoral comparisons as that between goods and services involves the necessity of facing interactions between these sectors squarely. We are going to investigate only certain aspects of firm behavior here and, hence, we shall take notice of such interactions only in passing. However, it may be essential to have a dynamic concept of behavior in order to discuss such interindustry relationships as occur in services that produce managerial assistance for use by goods producing firms. Such a dynamic behavior concept is our focus here.

It is rather easy to criticize existing measures of output and input in the services. We are all familiar with the fact that output is measured to equal input for some important subsectors (important both in terms of shares and in terms of employment trends), a procedure that effectively precludes measurement of technical structure. We are all aware of a problem of "quality"; we know that failure to adjust for changes in it leads to biases in conventional measures of productivity trends and we would probably agree that output measures for services are generally sorely afflicted by this problem. Many service outputs are highly intangible or qualitative in nature.

However, these are not directly the issues I want to consider; solutions to these measurement problems will be found in careful studies of individual industries such as those to come later in the present program.[7] *General* solutions are almost trivial as far as theory goes: one

a difficult "industry" to study for many reasons. Though it is a well-known cliche that this industry is "inefficient," the evidence is scant. Indeed, it appears that when a concentrated effort is made to measure output, this cliche is not substantiated. See [8] for such a rare attempt.

[6] This pattern has been observed by many economists and is the subject of a forthcoming book by Delehanty [5]. Being myself employed in the service sector, I must admit that I am biased against these findings of differential productivity trends from the beginning and I find the term "nonproduction worker" when applied to people doing what I do somewhat less than apt.

[7] Of course, to the extent that output is measured to equal input in certain key industries, a large theoretical problem exists, stemming from the fact that these industries interact with others through nonmarket channels. The inadequacy of our treatment of these cases is further testimony to our failure to study such channels.

adjusts for quality by reference to whatever quality indices one can come up with for the particular case. Economists may argue about direct imputations versus multiple regression methods. Nevertheless, we must agree that qualities do matter *and* that these are peculiar to each industry. What I want to consider, however, are general problems, ones that afflict any comparative study of the structure of production and, hence, ones that may be partially amenable to a general treatment.

What are the assumptions typically found in production studies, assumptions that might lead to errors in attributing productivity growth to one sector as opposed to another? The list is as familiar as the ten commandments even if the consequences of evil are less well established. Perfect competition, equilibrium, constant (internal) returns to scale, and the absence of externalities are usually assumed. The latter assumption carries implicitly with it the rather amorphous assumption that government behavior is neutral. One further assumption, seldom voiced but nevertheless implied, is that production takes place at a point; a rather unfortunate hypothesis since it immediately excludes locational patterns as a source of information.

The relaxation of almost any one of these typical assumptions is likely to have significance for intersectoral comparisons. I am going to consider one issue in detail: the problem of equilibrium and disequilibrium. This problem is pervasive in modern economic analysis; it has general significance. Furthermore, it seems plausible that the formulation of a theoretical framework for analysis of disequilibrium behavior is essential if we are ever going to understand the sort of phenomenon the service sector has exhibited in the last few decades.

The notions of "short-run" and "long-run" behavior are at least as old as Marshall's *Principles of Economics*.[8] Every undergraduate student of economics has seen the long-run average (marginal) cost curve and its corresponding short-run average (marginal) cost curves derived by holding one of the factors of production fixed. Every graduate student of economics has read Samuelson and knows about the envelope theorem.[9] On a rather different plane, virtually every economist has read dissident discussions about the motivation of the firm:

[8] The first edition came out in 1890.
[9] See [30], pp. 34–36.

maximum profits, growth or what have you.[10] These are often based on some empirical argument having to do with dynamic behavior, a class of phenomena for which there exists little or no theory of the firm.

In the light of these traditional acknowledgements of the problem of disequilibrium behavior, it is truly remarkable that theorists have done so little to develop a more general theory. Only recently has the itch of the aesthetic incongruity of the neoclassical (equilibrium) theory of the firm coupled with the fixed-weight lag schemes used by empirical workers produced attempts to generalize the theory.[11] Eisner and Strotz ([7], pp. 63–86) introduced a very simple model of the firm with a "cost of adjustment," which was then extended by Lucas [25]. Their purpose was to rationalize the distributed lag models.[12] The present paper is an attempt to indicate the more general applicability of a similar theoretical framework.

At this point one might well ask, "But what relevance does such a theory have for the services?" Aside from the spectacular growth in employment in the services that should suggest the possibility of disequilibrium, there is reason to believe that many of the services use relatively little physical capital.[13] More than in many other activities, growth in human resources is likely to be a dominant factor not only as evidenced in numbers but also in terms of internal investment; labor is likely to be the major overhead factor. The costs of this

10 An example is Penrose [29], pp. 26–30.

11 Nothing is quite so illustrative of the heights this schism between theory and practice has reached as the recent literature on investment. One observes the keen theoretical mind of a Jorgenson (e.g., [18], [19]). He elaborately constructs a theoretical model of the firm's *equilibrium* capital stock, using a patently dynamic tool of analysis (the calculus of variations) unnecessary for that task, and then, having obtained precisely the familiar set of *static* necessary conditions, he appends in an arbitrary fashion a distributed lag adjustment mechanism whose presence is dictated by nothing more than the econometric necessity of representing a dynamic process. If the lag structure is a constraint, why is it not part of the optimum problem? If it is merely a descriptive appendage, what is the role and relevance of the postulate of maximizing behavior in the first place?

12 I have criticized this rationalization somewhat ineffectually in [34] where I also considered a theory of rational excess capacity and the Keynesian marginal efficiency of investment function based on the adjustment cost notion. Notice should be made also of several recent papers by Lucas [23], [24], which study the dynamics of a competitive industry with constant returns to scale.

13 See Fuchs [11], p. 19.

process of expansion, internal through "learning-by-doing" and external through additional numbers, may result, if ignored, in a substantial error in measured technical relationships.[14] Furthermore, it would seem in the nature of many services that their product is tantamount to investment for the purchasing industry, though this is not accounted for as an increase in some quasi-fixed factor for the customer; this then biases his measured productivity upwards.[15]

Finally, there is an apparent shifting in service activities from manufacturing firms to specialized firms classified as services.[16] This sort of process of a widening "division of labor" that comes about through increases in "the extent of the market" has often been recognized but cannot be characterized analytically or measured empirically without a theory of dynamic behavior. That economies of scale are related to this phenomena is obvious.[17] But economies of scale may often be the more or less short-term evidence of the inflexibility of some factor. It is clear that the characteristic of fixity of any productive source is relative. A process of integration or disintegration of the above sort must come about as a result of the *relative* flexibility of various factors of production.[18] It follows that we cannot expect to evaluate the empirical significance of so-called economies of scale without a theory that consistently incorporates them; thus the succeeding analysis should be relevant to the generalization of one of the other traditional assumptions of productivity measurement, constant returns to scale.[19]

[14] In this context, the existing practice of speaking in one breath about professional or business services on the one hand and personal services on the other is incongruous. Given the radically different nature of the human resources used in these two kinds of "services," the generalizations one would expect to be able to make about the aggregate are rather limited.

[15] A case in point is advertising, which is recognized as a form of investment. See Arrow [2] and Nerlove [28].

[16] An account executive in a leading Chicago advertising enterprise informs me that this shift has taken place in this industry almost entirely since World War II.

[17] See Stigler [32] for a classic discussion of this process, using cost curves as the tool of analysis. That some dynamic process is behind his description is clear, but that aspect is left entirely implicit.

[18] The related process of merger is likewise based on a theoretical framework of sand without recognition of the relative flexibility of productive resources.

[19] In Kendrick's introduction to a previous Studies in Income and Wealth volume [20], he says (p. 9), "One suspects that we may in the end have to rely more on the hunches of a Stigler than on neat econometric solutions to the puzzle of distinguishing the effects of scale economies from those of technical progress." This statement is made in reference to Stigler's paper (pp. 47–63 in the same volume) in

I close this section by reminding the reader that I will provide only the most casual empirical foundation for what follows. In section II, I outline a general framework that can be evolved into a disequilibrium theory of the firm. We specialize greatly at the end of that section and section III is a derivation of the conditions necessary for the optimal behavior of a firm under these specializations plus a series of illustrations of this framework. Section IV concludes the paper.

II. A GENERAL STRUCTURE FOR THE THEORY OF ENTREPRENEURIAL DECISION

In this section I will present an analytical framework that, when specialized and simplified in given empirical contexts, can be evolved into a dynamic theory of the firm. I emphasize that we do not have here a complete, coherent theory of the firm, which would require elaborate empirical and theoretical work far beyond the scope of this paper.[20] However, I will specialize the analysis in the succeeding section to indicate the relevance of this framework to several particular problems.[21]

We take the firm to be a collection of existing resources together with a motivation to maximize its present value subject to a technical constraint. It is possible to justify this motive of present value maximization in a manner analogous to Irving Fisher's famous proof that maximization of consumer (stockholder) utility can be decomposed into two problems: (1) maximization of the present value of asset holdings and (2) maximization of utility subject to given present

which the author identifies several sources of measured productivity change: (1) changes in relative factor prices, (2) changes in industrial organization, (3) economies of scale and (4) technical change. Since, in the context of productivity measurement, no one knows what technical progress is, as is evidenced by the booming occupation of giving it a name, it seems unlikely we can reasonably speak of distinguishing it from something else. Nevertheless, we can analyze the first three problems Stigler mentions, if we have an adequate theoretical framework and a set of data rich enough to consider these problems.

20 Even less do we have here a model suitably applied directly to data at aggregate levels. We are operating under the assumption that our understanding of aggregates depends on our conceptualization of disaggregate behavior, in this case the behavior of the firm.

21 I think that a similar framework can be useful in the analysis of the behavior of economic organizations that do not maximize present value; they still are likely to face the kind of constraints represented by equation (4) below.

value.[22] This is relevant in that the succeeding structure is capable of generating many of the patterns of entrepreneurial behavior sometimes asserted to require "more general" criteria of behavior.[23]

The existing resources are designated by a vector K. Each element of K is a stock. These stocks yield a vector of service flows S, which may receive current wages W.[24] The resources may be accumulated by purchasing them externally at gross investment rates I, perhaps requiring some payment per unit, G. The accounting costs for a period are thus:

$$W'S + G'I \tag{1}$$

where a prime indicates vector transposition.

The firm sells a set of products Q currently, at prices P, generating a gross revenue: [25]

$$P'Q. \tag{2}$$

In order to compute the present value of the firm, we must, in general, specify market conditions and expectations. To simplify we shall consider the case of a price-taking firm.[26] The simplest expectations hypothesis is that of stationary expectations: prices are viewed as fixed at current levels. This is, of course, implausible for some purposes, but there is no obviously more plausible alternative.[27]

[22] See Fisher [9]. Hirschleifer [17] has considered the same and related problems recently. The assumption of a competitive capital market is generally needed.

[23] For example, Penrose [29], pp. 26–30. The issue is frequently clouded by the erroneous notion that some particular dividend policy is implied by maximization of "profits." There is nothing inconsistent between large retained earnings or growth of firms and the hypotheses of present value maximization and a competitive capital market, if we recognize a more general technical constraint than the "production function" of conventional theory.

[24] S and W are of the same dimension as K, as are G and I. P and Q below are of equal dimension.

[25] We ignore taxes and subsidies to simplify the exposition.

[26] This can be generalized in a number of directions; the case of given finitely elastic market relationships is easy to analyze even if these are shifting exogenously in time. A difficult generalization is to be found in dynamic market conditions that depend on the firm's behavior. See Graves and Telser [14] for an analysis of this kind of problem if cost curves are given.

[27] Any "certainty equivalent" type path of expected prices can, in principle, be used in the present formulation as long as present value is well-defined. The cost

We assume that the firm can borrow or lend at an interest rate r. Its present value is thus: [28]

$$V = \int_0^\infty [P'Q - W'S - G'I]e^{-rt}\, dt. \tag{3}$$

The present value is postulated to be maximized subject to a general technical constraint: [29]

$$F(Q, \dot{K}, S) = 0. \tag{4}$$

The constraint (4) has precisely the content of the "production function" of conventional theory for a given vector $\dot{K} = \dot{K}_0$, in particular in "equilibrium" where $\dot{K} = 0$. The present framework generalizes in such a way that a dependency between the production and expansion activities of the firm is allowed. Furthermore, (4) explicitly allows different resources to have different accumulation characteristics; some may be perfectly variable, others perfectly fixed and still others intermediate in this respect, variable in a restricted sense.

To specify the structure further we must relate service flows of the stocks to observable variables as well as saying something about "depreciation," the relationship linking $\dot{K}$ and I. We shall make two very special assumptions for present purposes, that the services of a given resource are proportional to its quantity and that depreciation is nil: [30]

$$\begin{aligned} S &= K, \\ \dot{K} &= I. \end{aligned} \tag{5}$$

These two assumptions are made here for no reason but expositional convenience. Generalizations of various kinds are possible. Deprecia-

in ease of manipulation of the model is, however, very large and analytical solution may be impossible. It should be observed that expectations are essentially irrelevant in models allowing perfect adjustment to existing conditions. They become relevant only if some such constraint as (4) is imposed.

[28] The time dependence of (Q, S, I, K) is ignored in the notation used here, though such dependence is implicit.

[29] A dot over a variable indicates a time derivative.

[30] The factor of proportionality between S and K can be taken to be unity without additional loss of generality.

tion in time can be postulated with no essential change in any analysis using this framework. A more interesting generalization is to introduce an intensity of use variable (possibly itself observed through some more variable input) that affects both the services actually realized from a stock and the depreciation (or perhaps appreciation through a learning process) incurred in the course of current operations. This kind of model can be used to generate a theory of rational excess capacity.[31] Most existing studies of production or investment treat excess capacity as if it were exogenous.[32] This is a most unsatisfactory state of affairs, since it sheds no light on the obviously mutual relationships between excess capacity, investment, unemployment or other relevant variables and excludes the investigation of a potentially important element in any dynamic theory of macroeconomics.[33] However, we will ignore all problems of that sort here.

At a given point in time ($t = 0$), the firm inherits a set of resources from the past: [34]

$$K(0) = K_0. \tag{6}$$

The formal problem facing the firm is thus:

$$\begin{aligned} &\text{Maximize} \int_0^\infty [P'Q - W'K - G'\dot{K}]e^{-rt}\, dt \\ &\{K(t): t > 0\} \\ &\qquad \text{Subject to:}^{35} \quad \text{(a)}\ F(Q, \dot{K}, K) = 0 \\ &\qquad\qquad\qquad\qquad \text{(b)}\ K(0) = K_0. \end{aligned} \tag{7}$$

This maximization is carried out by a choice of a path $K(t)$ for the firm's resources.

The innovation in the present approach is in the inclusion of the expansion rates $\dot{K}$ in the technical constraint. We view the firm as "producing" two sorts of "output": (a) that set of items (Q) which it

[31] See [34], pp. 10–36 for such a model.

[32] See Solow [31], Berglas [3] and Bourneuf [4], a more or less random sample.

[33] To say nothing of the possible simultaneity bias implied by this procedure.

[34] This existing stock of resources is thus data for the firm's decision. Whether or not this data is relevant depends on the character of (4).

[35] We ignore nonnegativity constraints on K, Q.

sells, including some possibly sold at zero prices (externalities) and (b) a set of *rates* of accumulation of productive assets to be used in the future.[36] We regard current investment as bearing directly on the current revenues obtainable from a particular combination of existing resources. A large class of propositions about firm behavior can be obtained by suitable restrictions on the function F; in particular, the hypothesis of equilibrium itself ($\dot{K} = 0$) could conceivably be evaluated empirically within this kind of framework, an obvious impossibility within any model *assuming* equilibrium.

We can illustrate the proposed formal framework with a few examples. Consider a perfectly variable input K_i. It is represented above by recognizing that $(\partial F/\partial \dot{K}_i) \equiv 0$, that is $\dot{K}_i$ does not enter the technical relation F. It may or may not be true that $G_i = 0$. Examples would be purchases of raw materials consumed within the accounting period, some kinds of labor (e.g., Kelly Girls) and some kinds of rental equipment, those not involving contractual arrangements that extend some commitment to rent outside the accounting period.

On the other extreme, consider a perfectly fixed factor K_j. If the vector K is a complete enumeration of all resources involved in the firm, logic should require the following constant returns condition to hold:

$$F(\lambda Q, 0, \lambda K) = 0 \text{ for all } \lambda > 0 \tag{8}$$

where $\dot{K} = 0$. If K_j is perfectly fixed, it can be eliminated from the vector K and we should expect some pattern of increasing or decreasing returns or both to occur for the included set of resources. The price G_j is irrelevant as $\dot{K}_j \equiv 0$, but W_j (the rent to K_j) may or may not be zero.[37]

In general, a resource K_1 will be neither perfectly fixed nor perfectly variable, though this depends on the period. In that case, $\dot{K}_1$

[36] Note that the firm is not "producing" the productive assets, rather their rates of accumulation, which have purchase prices (G) but, also, internal costs in foregone sales. Expansion must be planned and managed and, hence, diverts resources from the production-for-current-sale activities of the firm.

[37] Of course, it may be impossible to identify K_j and we may not want to call it a resource in that case. In that case, (8) is not logically implied.

will appear in F, indicating that its accumulation has some impact on "production" activities of the firm and G_1 and W_1 may or may not be zero. For physical capital goods $G_1 > 0$ and there may be a well-defined unit operating cost $W_1 > 0$ for the item. Human capital will typically earn a current wage W_1 and, in addition, there may be a market price of accumulation such as the costs of external training of existing human resources or the price of the services of an employment agency if this personnel operation is external. An internal on-the-job training program or an internal personnel department will have effects on current productive activities and these effects are represented in the function F by the inclusion of $\dot{K}_1$ as an argument.

There may be resources relevant to a given process F that are external to the firm and that enter F in stock or flow form but that do not enter the net revenue function. These would typically have zero flow and stock prices: the technical knowledge produced by a government research and development operation, for example. Hybrids may occur in the externality case: a firm may not pay for the accumulation of some stock of external goods, but may pay a rental for the use of the stock. The price W_1 would then be positive while $G_1 = 0$. For example, roads may be built that imply no recognized accumulation cost to trucking firms, but tolls may be charged that imply a current operating cost; this becomes a variable input if we recognize the possibility of variable use levels.

Before turning to the very special case we shall consider in the rest of this paper, comment must be made on measurement problems. In the case of variable inputs, measurement is likely to involve observation of market data which must then be decomposed into prices and quantities. The problems that may arise here are familiar. The same statement applies to the firm's outputs, Q. A major qualification as to externalities must be made; observation of market data may be inadequate in studying any firm whose technical conditions depend on the level of certain chemicals in the air, if that firm is located on the South Side of Chicago. Likewise, errors in the opposite direction will accompany measurement of the output of the "polluters." Is there reason to expect external effects of "services" to be positive as opposed to the negative externalities of many "goods" industries? In certain leading cases, it seems plausible that much of output should be viewed as an externality: government. Productivity for this industry will never be understood without making use of this fact.[38]

As for the measurement of imperfectly variable resources, many problems are the same here as in equilibrium theory, but some are added. Problems of distinguishing net as opposed to gross investment are well-known.[39] Even if we can, through observation of market transactions in capital goods, measure gross physical investment accurately, we must know how depreciation occurs if we are to compute a stock of even homogeneous units.[40] This problem may be somewhat clarified by an explicit theory of firm expansion; it will, however, probably remain a thorn in the flesh of empirical workers. To the extent that there is a current wage (W) paid to an imperfectly flexible factor, this may serve as a further means of measuring the stock of that factor, especially if a theory linking services and stock can be constructed along with a theory of the evolution of internal resources. It is, however, an error of fundamental importance to assume that, because a factor is paid a current wage (i.e., it is rented), it is perfectly flexible *from the viewpoint of the firm*. Highly skilled human capital frequently is rented under a contract binding on the firm and it is costly to contract for more of such capital either externally or internally. Failure to recognize this characteristic of human capital can produce most of the empirical anomalies in services by the sheer dint of the relative labor intensity of this rapidly expanding sector.

At this point we shall proceed to impose a particular set of very strong restrictions on the function F for purposes of exposition in the rest of this paper. These restrictions are only convenient for present purposes and should not be confused with the general framework. In general, the tools already developed in economics to deal with general

[38] This is the meaning of the use of R and D expenditures in production functions for industries in the private sector. For example, see Griliches [15]. This approach has not been carried very far as yet.

[39] See Tice [33] for an interesting study of the relevance of these problems.

[40] That even the measurement of real gross investment in physical resources is currently in substantial error is discussed and evaluated by Griliches and Jorgenson [16]. If errors in measuring the rate of accumulation of physical capital are inducing large measured rates of productivity change in the "goods" sector, the fact of the relative labor intensity of "services" may explain much of the observed differential in productivity trends.

functions can be applied to simplify F given, of course, that F satisfies conditions necessary and sufficient for (7) to have a solution.[41]

We are going to assume that there is one output, Q, which is sold at the price P, one perfectly variable input, L, rented at a wage, W, one imperfectly variable factor, K, which receives a current wage, R, for its services and is accumulated at a rate $\dot{K}$ (there is no depreciation), the external price of this investment good being G.[42] Furthermore, we are going to assume that F is characterized by the very restricted form:

$$Q + C(\dot{K}) - F(K, L) = 0, \tag{9}$$

where the notation F is here taking on a new definition.[43] The term $F(K, L)$ is analogous to the static production function and, by observing (8), we see that the full production function of (8) has constant, decreasing or increasing returns to scale to the extent that $F(K, L)$ of (9) has such a configuration, if we require $C(0) = 0$. It is convenient to assume that C and F have three continuous derivatives and it is useful to have positive and diminishing marginal products:

$$\begin{aligned} &F_K(K, L), F_L(K, L) > 0 \\ &F_{KK}(K, L), F_{LL}(K, L) < 0. \end{aligned} \tag{10}$$

Further restrictions will be discussed in section III.

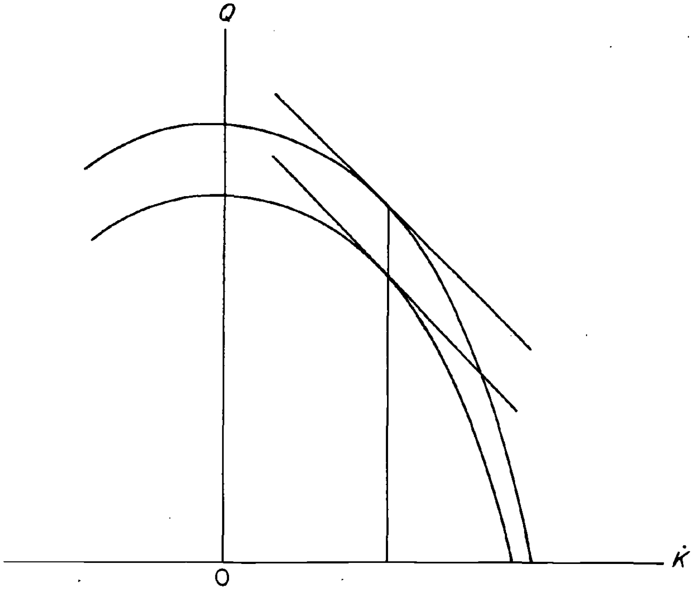
The restrictions we have imposed in (9) eliminate certain kinds of substitution that may be empirically relevant. The marginal rate of

41 In particular, various kinds of independence of the arguments of F as considered in the literature on separability would be useful. See Goldman and Uzawa [13] for a summary. Certain recent work on multiproduct production functions should also be relevant for the problem of selecting functional forms for empirical work. See Mundlak [26], and Mundlak and Razin [27].

42 There is no necessary connection between our symbols (K, L) and "capital" and "labor" as conventionally understood. Our K is merely a quasi-fixed factor while L is a perfectly variable one. There is no necessary suggestion that human factors are more variable than nonhuman ones.

43 One observation about the manner in which service outputs are often measured is clear. If (9) is an accurate representation of the technical constraint, then the measurement of output by inputs will not only preclude the measurement of productivity change. It will tend to give a grossly misleading impression of output trends or cycles to the extent that $\dot{K}$ behaves perversely, since the implicit assumption is that $C(\dot{K}) \equiv 0$.

FIGURE I



substitution between the output Q and expansion $\dot{K}$ is independent of both the existing resources (K and L) and the level of output Q. If we allowed for depreciation (or appreciation through learning) of K, the former assumption would be relaxed somewhat since such depreciation would depend on the resource K. The independence of the marginal rate of substitution between Q and $\dot{K}$ of Q is represented in Figure I for varying levels of $F(K, L)$. The slope of this "production possibility frontier" is merely:

$$\left.\frac{\partial Q}{\partial \dot{K}}\right|_F = -C'(\dot{K}).$$

We have drawn the figure strictly concave, though it need not be.

III. THE CHARACTER OF OPTIMAL BEHAVIOR

Under the restrictions we have imposed, the firm faces the following problem:

$$\left\{\begin{array}{ll} \text{Maximize} \displaystyle\int_0^\infty [PQ - WL - RK - G\dot{K}]e^{-rt}\,dt & \\ \{L(t), K(t)\} & \\ \text{Subject to:} \quad \text{(a)}\ Q + C(\dot{K}) - F(K, L) = 0, & \\ \qquad\qquad\qquad \text{(b)}\ K(0) = K_0. & \end{array}\right. \tag{11}$$

Substitution of the constraint (11a) into the integrand reduces the problem to the following:

$$\underset{\{L(t),\,K(t)\}}{\text{Maximize}} \int_0^\infty [PF(K, L) - WL - RK - G\dot{K} - PC(\dot{K})]e^{-rt}\,dt. \tag{12}$$

If this problem has a solution, it must satisfy the following necessary conditions:[44]

$$\begin{aligned} &\text{(a)}\ PF_L(K, L) = W \\ &\text{(b)}\ \ddot{K} = \frac{r[G + PC'(\dot{K})] + R - PF_K(K, L)}{PC''(\dot{K})} \\ &\text{(c)}\ \lim_{t\to\infty} [G + PC'(\dot{K})]e^{-rt} = 0 \\ &\text{(d)}\ C''(\dot{K}) \geqslant 0. \end{aligned} \tag{13}$$

The interpretation of (13a) is familiar: a perfectly variable input will be used in such a quantity as to equate the value of its marginal product PF_L to its wage, W. To interpret (13b), rearrange it to read:

$$\frac{d}{dt}[G + PC'(\dot{K})] = r[G + PC'(\dot{K})] + R - PF_K(K, L). \tag{14}$$

This equation can be integrated to yield:

$$G + PC'(\dot{K}) = \int_t^\infty [PF_K(K, L) - R]e^{-r(\tau - t)}\,d\tau, \tag{15}$$

which must hold for each future time, t, the left side being evaluated as of t. The right hand side is the marginal value at time t of invest-

[44] In the terminology of the calculus of variations, conditions (13a, b) are Euler equations, (13c) is a transversality condition and (13d) is a Legendre condition; all are necessary, though not sufficient, for maximum present value.

ment projected for t: the discounted sum of later net values of marginal products of the quasi-fixed resource, K. The left hand side is the contemporaneous marginal cost of investment of time t, which is composed of two parts, any unit market cost G that may exist and the marginal value of real product foregone as a consequence of expansion at the rate $\dot{K}$. Equations (15) or (13b) require that the marginal cost of investment be equated to its marginal value. The *present* ($t = 0$) value of investment tó be carried out at time t is equal to $[G + PC'(\dot{K})]e^{-rt}$ under optimal conditions. Equation (13c) requires that this present value approach zero as the date, t, of the investment recedes into the infinite future. One can interpret this latter condition heuristically by thinking of the finite horizon analog; we should not attach any value to investment carried out at the horizon, because we are systematically disregarding any returns thereafter that could contribute to such value. Condition (13d) requires the production possibility frontier of Figure I to be concave in a neighborhood of the solution at each point of time; it is in analogy to second-order conditions in static theory. To eliminate certain obvious nonuniqueness problems, we shall assume $C''(\dot{K}) > 0$ for all $\dot{K}$.

There are several patterns that may characterize the firm's dynamic behavior, depending on the character of $F(K, L)$. Consider first the case of decreasing returns to scale:

$$H(K, L) = F_{KK}F_{LL} - F_{KL}^2 > 0. \qquad (16)$$

We can construct a phase diagram in the $(K, \dot{K})$ plane to represent geometrically the course the firm takes. Consider the locus of points for which $\dot{K} = 0$ (in equation 13b):

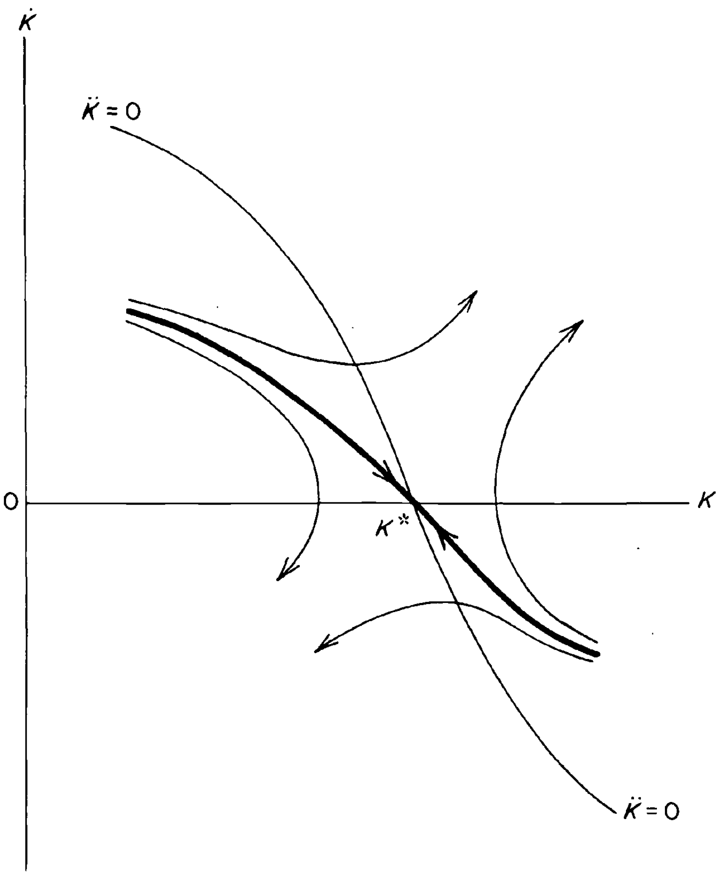
$$r[G + PC'(\dot{K})] + R = PF_K(K, L). \qquad (17)$$

Differentiate totally with respect to K (i.e., using 13a) to obtain:

$$rPC''(\dot{K}) \left.\frac{\partial \dot{K}}{\partial K}\right|_{\dot{K}=0} = \frac{PH(K, w)}{F_{LL}(K, w)} < 0 \qquad (18)$$

where $w = W/P$ is the real wage of the flexible factor. This locus has negative slope as shown in Figure II. The phase pattern shown there

FIGURE II



is then discernible from (13b).[45] There is a saddle point path (shown by heavy arrow), which converges to a (unique) equilibrium scale K^* defined by:

[45] Consider a given K and the $\dot{K}$ corresponding to it on the $\ddot{K} = 0$ locus. Because $C''(\dot{K}) > 0$, if we let $\dot{K}$ rise, the numerator of (13b) becomes positive and vice versa if we lower $\dot{K}$. Thus $\ddot{K} \gtrless 0$ as $\dot{K}$ lies above or below the $\ddot{K} = 0$ locus. Points above this locus are then characterized by a northward motion, points below the locus are characterized by a southward motion. Any point in the first quadrant is characterized by $\dot{K} > 0$ and, hence, by an eastward motion and vice versa for any point in the fourth quadrant.

$$\begin{aligned} &\text{(a)}\ PF_L(K^*, L^*) = W \\ &\text{(b)}\ PF_K(K^*, L^*) = R + r[G + PC'(0)]. \end{aligned} \tag{19}$$

Clearly this path satisfies (13c) and it is possible to show that it is unique in this respect.[46]

We pause to make a few comments on this case. It is well known that the product is not exhausted for a firm in equilibrium under diminishing returns to scale and that this is a source of error in conventional productivity measurements that use accounting income shares as measures of productive shares. Another error is introduced even if accounting shares are not used, if the term $C(\dot{K})$ is not recognized as offsetting potential output (F) in disequilibrium. It may be that $C(0) = 0$, though we seldom observe $\dot{K} = 0$ for physical capital accumulation at the levels of aggregation we are accustomed to studying.[47] This means that most firms may be seldom, if ever, in a position like K^*, but may usually be somewhere on the branch of the optimal path to the left of K^*. *The more rapidly they are expanding, the more understated will be estimates of their potential output* (F), because the larger will be the ignored $C(\dot{K})$.[48]

The preceding analysis is incomplete. A significant possibility is that of constant returns to scale. This case is rather simple and is shown in Figure III. There is a fixed rate of accumulation $\hat{\dot{K}}$ given by:

$$r[G + PC'(\hat{\dot{K}})] + R = PF_K(K, L) \tag{20}$$

where $F_K(K, L)$ is already fixed in (13a) since marginal products depend only on the ratio K/L. This rate of accumulation is independent of the existing scale and obviously satisfies (13c). Again, the light arrows indicate suboptimal paths.

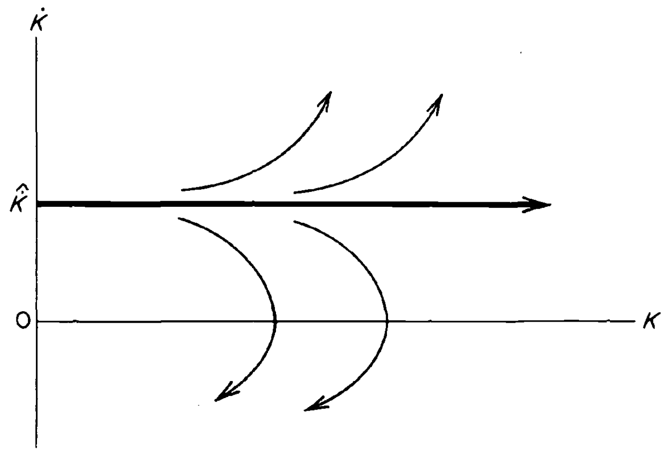
There are apparently a number of things that can be wrong in an analysis of technical conditions that fails to consider this kind of behavior. Note that we have a kind of equilibrium here. In a growing

[46] The proof is given in [34], pp. 76–77.

[47] Of course, we never observe net investment $\dot{K}$ anyway, only gross. But I know of few *estimates* of $\dot{K}$ that are not positive. This comment is particularly plausible if we remember that K here includes all imperfectly variable resources.

[48] This assumes $C' > 0$ for $\dot{K} > 0$ which seems plausible.

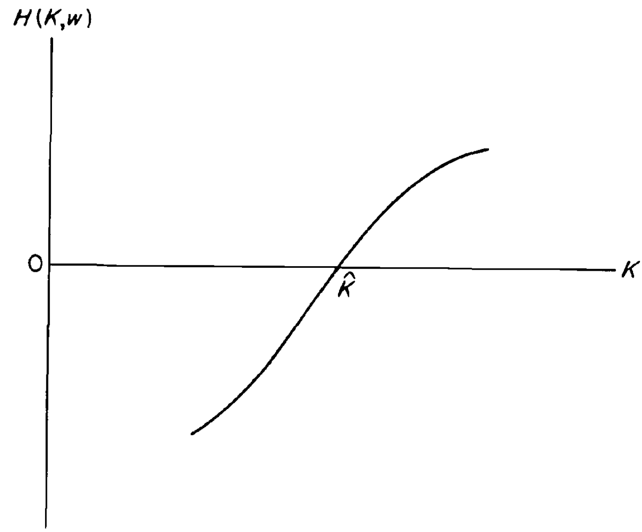
FIGURE III



economy, an industry can be made up of firms all of whom are growing. I venture to suppose that this is even typical. If all face the same relative price for variable inputs, they will all exhibit the same marginal productivity of capital if they enjoy constant returns to scale, but if we attempt to measure this marginal product by a ratio of income net of wages paid to the variable factor to some stock, the result in general confounds the effects of the expansion process with the "productivity" of the resource. This is true even if the stock is measured perfectly. The meaning of exhausting the product is unclear in a dynamic context even if there is constant returns to scale in the static production process.

We now turn to a further extension. What is the dynamic pattern of a firm facing a U-shaped long-run average cost curve? That is, suppose that $H(K, w)$ has a pattern as in Figure IV for a given real wage w of the variable factor. There are increasing returns to scale for $K < \hat{K}$ and decreasing returns to $K > \hat{K}$. We shall assume the monotonicity of H as in Figure IV. It is clear from (18) that the locus of points in the $(K, \dot{K})$ plane for which $\ddot{K} = 0$ will have an inverted U shape as in Figure V, which is drawn with this locus meeting the $\dot{K}$ axis at a positive level. The unique optimal path is shown with heavy arrows; the others are suboptimal. There is a particular optimal level of investment $\dot{K}_0{}^*$ which characterizes entrants to the industry. Once entered,

FIGURE IV



the optimal level of investment rises as the stock rises until a critical maximum is reached at $\tilde{K}$ and then recedes as the stock gets even larger; the stationary solution K^* is reached in the limit.

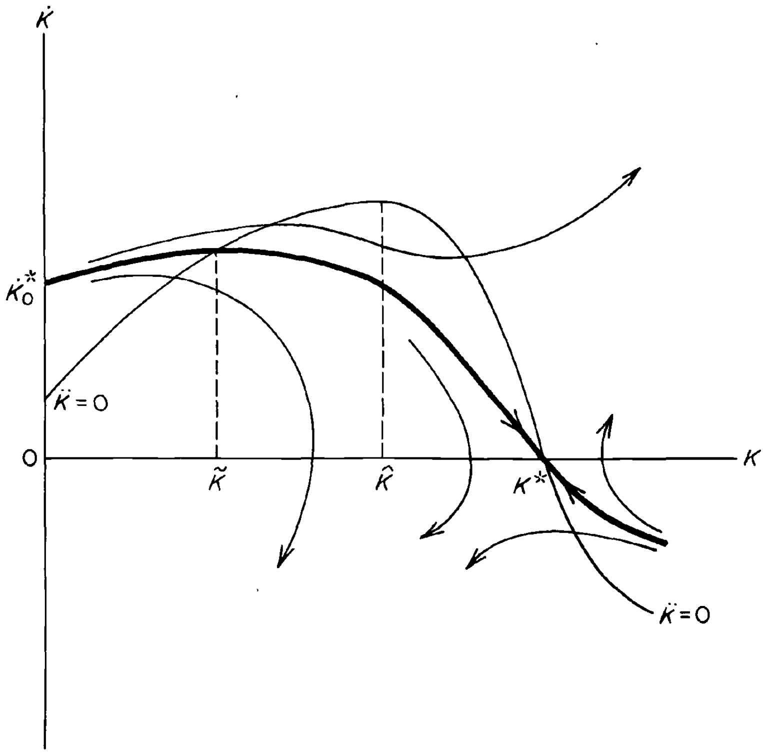
One use for this sort of model is to analyze conditions for entry and exit. As an industry develops, the level of the $\ddot{K} = 0$ locus may fall. As can be seen in (17) this will occur as P declines under certain circumstances or as R, G or r rises. For whatever reason external to the firm, the $\ddot{K} = 0$ locus may intersect the $\dot{K} = 0$ line (the horizontal axis) at another equilibrium point, $\bar{K}^*$, as in Figures VI and VII. This equilibrium cannot be a saddle point. To consider what configuration it takes, we utilize a linear approximation of (13b) in a neighborhood of $\bar{K}^*$:

$$\ddot{K} \simeq r\dot{K} - \frac{H(\bar{K}^*, w)}{C''(0)F_{LL}(\bar{K}^*, w)}(K - \bar{K}^*) \tag{21}$$

The roots of this equation are: [49]

[49] Note that at K^*, rather than $\bar{K}^*$, $H > 0$, so that the roots are real and of opposite sign, indicating the saddle point about K^* already observed above in Figure II.

FIGURE V



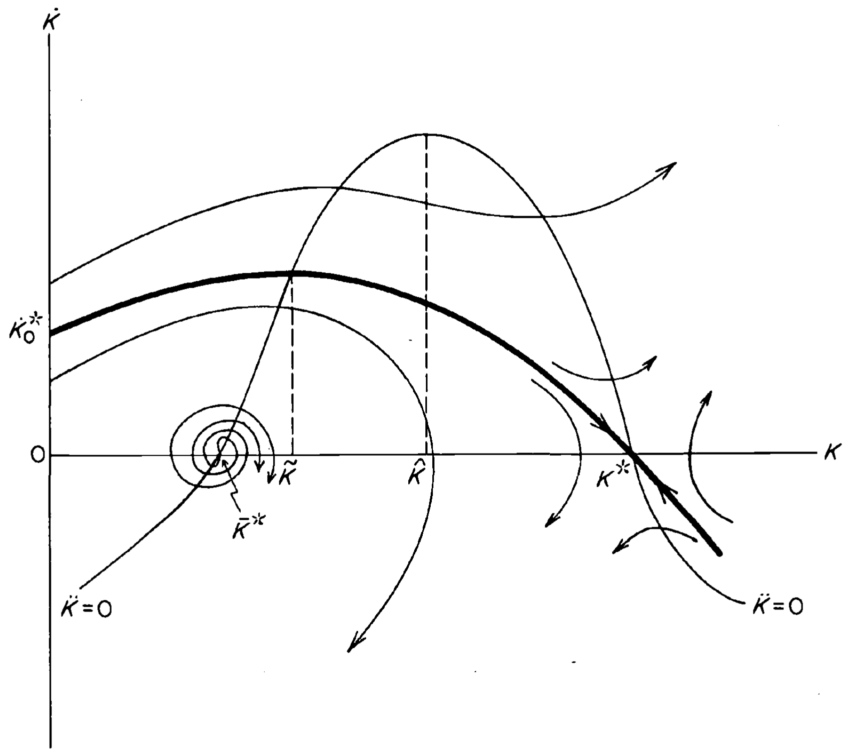
$$\psi = \frac{r}{2} \pm \sqrt{\left(\frac{r}{2}\right)^2 - \frac{H(\bar{K}^*, w)}{C''(0)F_{LL}(\bar{K}^*, w)}}. \tag{22}$$

Since $H < 0$ at $\bar{K}^*$, the discriminant of the roots may or may not be positive, but in either case the real parts will be positive, indicating that $\bar{K}^*$ is an unstable equilibrium. If $H(K, w)$ has the configuration in Figure IV (i.e., monotonic increasing), then there is some scale $\bar{K}(w)$ at which: [50]

$$\left(\frac{r}{2}\right)^2 = \frac{H[\bar{K}(w), w]}{C''(0)F_{LL}[K(w), w]}. \tag{23}$$

[50] For simplicity, I assume that $\bar{K}(w)$ is unique, given w, which implies a restriction on F_{LL} as well as H. When $\bar{K}^* = \bar{K}(w)$, the equilibrium is degenerate.

FIGURE VI



If $\bar{K}^* < \bar{K}$, then the roots ψ are complex and $\bar{K}^*$ is an unstable focus as in Figure VI. If $\bar{K}^* > \bar{K}$, the roots are real and positive so that $\bar{K}^*$ is an unstable node as in Figure VII. In the first case, if an optimal solution exists for all scales, there will be entry at a level of investment $\dot{K}_0^*$.[51] In the second case, $\bar{K}^*$ is a critical minimum scale that must be exceeded for the firm to expand; entry will not take place and firms of scale $K < \bar{K}^*$ will leave the industry as indicated by the lighter arrows proceeding left from $\bar{K}^*$.[52] This discussion indicates that a necessary condition for entry will be:

[51] $\dot{K}_0^* > 0$ clearly. Sufficient conditions for these two cases have not been derived.
[52] There is an ambiguity in the present formulation in that it is unable to specify which of these paths is optimal for "leaving the industry." This is due to the fact that we have only specified alternative *financial* opportunities; if the character of alternative productive activities, which must exist for r to be sufficiently high to

FIGURE VII

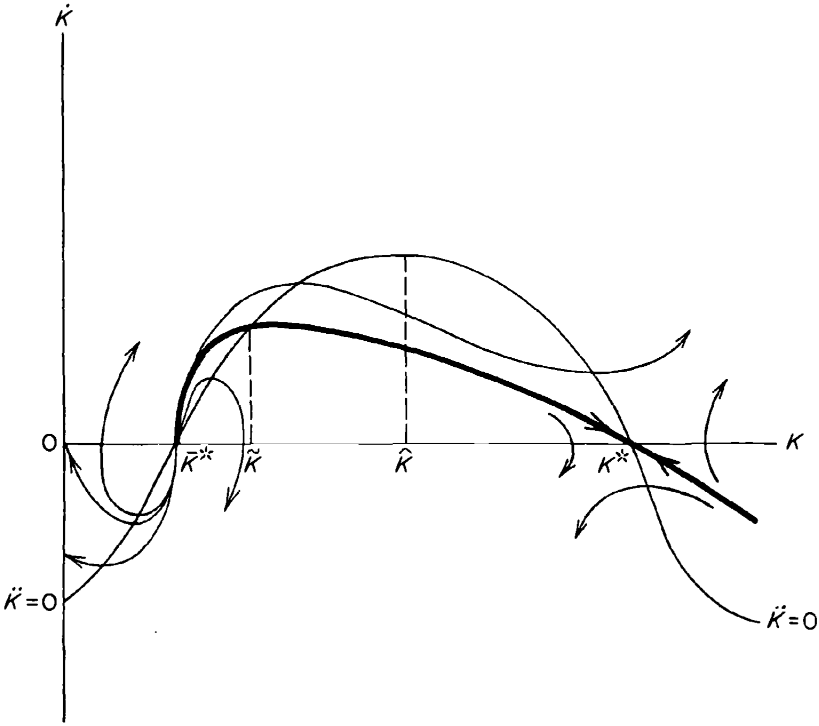
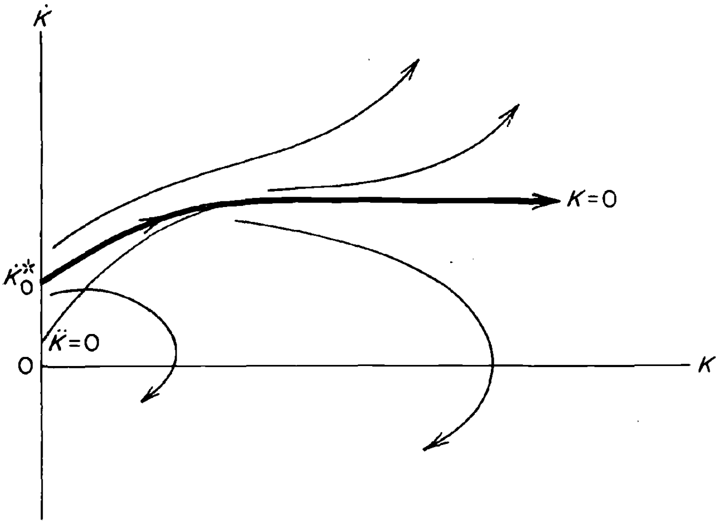


FIGURE VIII



$$\left(\frac{r}{2}\right)^2 < \frac{H[\bar{K}^*(w), w]}{C''(0)F_{LL}[\bar{K}^*(w), w]} \text{ if } \bar{K}^*(w) > 0 \tag{24}$$

while the reverse inequality will be sufficient to rule out entry and to possibly produce some exits (firms characterized by $K < \bar{K}^*$), all expansion taking place in existing firms.

It may be the case that a production process is characterized by a region of increasing returns to scale for small firms followed by a substantial range of constant returns. Figure VIII indicates the dynamic pattern in the case in which the $\dot{K} = 0$ locus reaches the K axis at a positive value (as in Figure V); analogs to Figures VI and VII are apparent.

The discussion has been aimed at demonstrating the usefulness of the kind of disequilibrium theory I outlined in section II. We see that it can be used in the investigation of several theoretical matters: the pattern of optimal investment, and the conditions under which entry or exit will occur. A similar framework can help shed light on merger processes and other dynamic patterns of industrial organization. It should be apparent in the foregoing that the dynamic pattern of firm behavior can be substantially dependent on the returns-to-scale characteristics the firm faces. This is well-recognized in existing economic literature; one innovation in the kind of theory we have proposed is that this dependence can be consistently incorporated in an analytical framework for which parametric representations suitable for empirical work can be designed, i.e., if we choose functional forms for $F(K, L)$ and $C(\dot{K})$ and identify the factors in a meaningful fashion, we can construct an operational econometric model of the firm facing increasing returns. We have considered a model with only two resources, but our allowance for economies of scale indicates the possible existence of some further factor fixed to the firm but not explicitly included in our analysis. Such factors, external or internal, need not be *perfectly* fixed. If internal, the generalization can be directly made and our problem increases in dimensionality. The more interesting

produce this situation, are specified, then a choice of these paths can be made. There will be an optimal pattern of shifting resources from one industry to another, depending on the dynamic constraints of the other industry as well as the present one.

problem, however, is that of external factors. How do they come about? How can they be characterized within the above framework?

We do not answer these questions but only illustrate their relevance with a few examples. As a first example, we should expect that space economies (or diseconomies) would be a pervasive factor in economic development. As space is more intensively utilized, the spatial pattern of production will change; this pattern will interact with developments in transportation and communication and the implications will differ for different products. That is, transport costs are an "adjustment cost" in many production processes, but they are more relevant for some products than others. It is recognized that technical progress in transport will allow the exploitation of economies of scale in agriculture and the fact that a government road building project will have external effects on *both* transportation and agriculture is well known. A production study constraining all processes involved to constant returns to scale will tend to overstate "productivity change" in the latter case for both transport and agriculture. But what sort of differential effects on other industries can we expect? Almost by definition, many services have peculiar transport characteristics. One cannot ship a haircut! The new road is, therefore, likely to have a minimal impact on the barbering industry. Other services may be vitally affected by a change in technology in another industry or by changes in the quantity of certain government services. For example, what was the role of improved communications, data processing technology and Federal Reserve services to member banks in the development of banking in the last few decades where we observe the rapid expansion of employment typical of the services. Aside from the necessity of developing better measures of "output" for banking, a requirement for the efficient study of almost all services, the study of this development would seem to require a dynamic conceptual framework incorporating these external patterns of change.

As external factors move over time, the shape and height of the dynamic paths shown in various figures above will change. That is, the character of the returns to scale configuration may shift. Observation of these external factors in conjunction with observation of changes in firm controlled variables will supply additional information about production structure entirely ignored in an equilibrium framework which attributes any change in, say, output to a change in some productive resource or in an unexplained residual. The residual sops up "our ignorance" as Abramovitz [1] has put it. Our ignorance is pervasive. Not only does it include imperfect measurements of many kinds; it incorporates our surpassingly imperfect conceptualization of dynamic processes.

IV. CONCLUSION

In conclusion, I want to describe what I regard as possibly fruitful directions for future research on the problems I have suggested above. Though the discussion will isolate empirical and theoretical problems, I take it as obvious that these must be considered jointly. Whitehead has said,

> All the world over and at all times there have been practical men, absorbed in "irreducible and stubborn facts": all the world over and at all times there have been men of philosophic temperament who have been absorbed in the weaving of general principles. It is this union of passionate interest in the detailed facts with equal devotion to abstract generalization which forms the novelty in our present society. ([35], p. 4.)

It is far from clear that this novel union is yet pervasive in economics.

In the realm of theory, the kind of framework presented in Section II promises a method of generalizing the theory of the firm to integrate investment and production theory, to describe intertemporal patterns of industrial organization and to describe general dynamic response patterns now represented in ad hoc fashion by distributed lags. Theoretical work providing an axiomatic basis for particular restrictions on the function F in (4) is needed. We must make a parallel effort to represent, and measure, the distinction between potential and actual services yielded by a given resource. There is, further, a great deal to be done in formulating econometric methods for use in this kind of model. Stochastic analogs of the necessary conditions (13) must be devised and there is a pregnant question of how to treat an end point condition like (13c). Related to this is the need for an analysis of intertemporal patterns in the size distribution of firms. Such an analysis can be based on our approach, but must go on to investigate the aggregative aspects of the problem.

As for measurement, we face all of the known problems of measuring prices and quantities involved in market transactions, of allowing for changes in qualities and of computing stock estimates for accumulated resources. This latter difficulty leads into the problem of identifying resources by their accumulation and depreciation characteristics. The traditional distinction between human and physical resources is too gross as a classificatory scheme for economic resources as, of course, has been recognized for some time. We have tried to measure skill levels, education and other demographic determinants of the productive quality of human capital. And the problem of "overhead labor" has been the subject of much recent research. Nevertheless, a substantial problem exists of specifying and measuring the growth or decay in the productive capacity of a given unit of any kind of resource and part of that problem is that market relationships inadequately reflect such processes for imperfectly flexible resources. Specialized nonmarket data on individual establishments or firms should be more systematically collected and a greater effort to analyze microeconomic phenomenon with the use of such microeconomic data should be made.

In the latter connection it seems to me that the national income accounting framework is peculiarly unsuited for the study of production structure. A convention like that of measuring output by "value-added," which may have a justification in computations aimed at measuring total output for an economy with little trade, if carried over to the measurement of the output of particular subaggregates, is unfounded on any valid theoretical principle and will undoubtedly lead to erroneous inferences in many cases. This is true even if all the rigid assumptions of the usual productivity analysis are made; if disequilibrium is allowed, it is virtually impossible to say what kinds of distortions value-added measures imply. Aside from the fact that the distinction between current and capital account is often fuzzy, the notion that "internal" resources contribute to output in an essential fashion while "purchases" do so only inessentially is based upon considerations that need have no relationship to the relevant technical or organizational structure.[53] If Q is output and M is purchases, both in

[53] This may be particularly relevant to comparisons between "services" and "goods," since the former are likely to be characterized by a smaller share of "purchases" to true output than the latter, a fact that will indicate smaller errors

real terms, the use of *Q-M* as an output measure implies an obvious loss of information. (*Q-M* is consistent with an infinite number of *Q*, *M* pairs.) We could never do worse by treating *Q* as output and *M* as an input and could often do better.

As for the use of factor shares computed from income accounting data, if there is disequilibrium, then these need not reflect marginal productivities, an observation made repeatedly by critics of neoclassical growth theory. Not only does this raise questions about the relevance of the conventional theory of income distribution, but it suggests that errors in measuring production structure will be induced by such procedures. How significant such errors may be cannot be asserted without empirical work within a more general conceptual framework.

There are then massive conceptual and empirical difficulties in the application of the framework proposed in this paper. This is, however, in the nature of any new approach. We recognize a set of problems: intertemporal patterns of measured productivity, industrial organization and investment behavior as well as problems of measuring scale economies. I have presented a conceptual framework in Section II that may be useful in the analysis of these problems as is suggested in Section III where a few examples are presented. If we were all constrained by the necessity of computing a number tomorrow to be decreed as a "guideline," this would be an utterly fruitless exercise. Given a somewhat longer horizon, we may find such an approach useful in the analysis of the above dynamic problems and in the analysis of others as yet imperfectly conceived.

REFERENCES

1. Abramovitz, M., "Resource and Output Trends in the United States since 1870," *American Economic Review*, vol. 46 (1956), pp. 5–23, reprinted as NBER Occasional Paper 52, New York, 1956.
2. Arrow, K. J., "Optimal Capital Adjustment," in K. Arrow, S. Karlin, and H. Scarf (eds.), *Studies in Applied Probability and Management Science*, Stanford, 1962.

in measured productivity change induced by the value-added convention for the services as long as the usual assumptions of equilibrium, constant returns and competition hold.

3. Berglas, E., "Investment and Technological Change," *Journal of Political Economy,* vol. 73 (1965), pp. 173–180.
4. Bourneuf, A., "Manufacturing Investment, Excess Capacity, and the Rate of Growth of Output," *American Economic Review,* vol. 54 (1964), pp. 607–625.
5. Delehanty, G. E., *Nonproduction Workers in U.S. Manufacturing,* Amsterdam, forthcoming.
6. Denison, E. F., *The Sources of Economic Growth in the United States and the Alternatives Before Us,* New York, 1962.
7. Eisner, R., and Strotz, R. H., "Determinants of Business Investment," Research Study Two in *Impacts of Monetary Policy,* Englewood Cliffs, 1963.
8. Executive Office of the President, Bureau of the Budget, *Measuring Productivity of Federal Government Organizations,* Washington, D.C., 1964.
9. Fisher, I., *The Theory of Interest.* Reprints of Economic Classics. New York, 1965.
10. Fuchs, V. R., *Productivity Trends in the Goods and Service Sectors, 1929–61: A Preliminary Survey,* Occasional Paper No. 89, New York, NBER, 1964.
11. ———, "The Growing Importance of the Service Industries," *The Journal of Business,* October 1965, pp. 344–373, reprinted as Occasional Paper 96, New York, NBER, 1965.
12. ———, and Wilburn, J. A., *Productivity Differences Within the Service Sector.* Occasional Paper 102, New York, NBER, 1967.
13. Goldman, S. M., and Uzawa, H., "A Note on Separability in Demand Analysis," *Econometrica,* vol. 32 (1964), pp. 387–398.
14. Graves, R. L., and Telser, L. G., "Optimal Output and Price Policies for General Distributed Lag Demand Equations," *Econometrica,* forthcoming.
15. Griliches, A., "Research Expenditures, Education, and the Aggregate Agricultural Production Function," *American Economic Review,* vol. 54 (1964), pp. 961–974.
16. ———, and Jorgenson, D. W., "Sources of Measured Productivity Change: Capital Input," *American Economic Review,* vol. 56 (1966), pp. 50–61.
17. Hirschleifer, J., "On the Theory of Optimal Investment Decision," *Journal of Political Economy,* vol. 66 (1958).
18. Jorgenson, D. W., "Anticipations and Investment Behavior," in *The Brookings Quarterly Econometric Model of the United States,* Chicago, 1965.
19. ———, "Capital Theory and Investment Behavior," *American Economic Review,* vol. 53 (1963), pp. 247–259.
20. Kendrick, J. W. (ed.), *Output, Input and Productivity Measurement,* Studies in Income and Wealth, vol. 25, Princeton for NBER, 1961.

21. Kuznets, S., *National Income and Its Composition, 1919–1938,* New York, 1941.
22. Lengellé, M., *The Growing Importance of the Service Sector in Member Countries,* Paris, O.E.C.D., 1966.
23. Lucas, R. E., "Adjustment Costs and the Theory Supply," *Journal of Political Economy,* vol. 75 (1967), pp. 321–334.
24. ———, "Optimal Investment with Rational Expectations," unpublished Carnegie Tech research memorandum.
25. ———, "Optimal Investment Policy and the Flexible Accelerator," *International Economic Review,* vol. 8 (1967), pp. 78–85.
26. Mundlak, Y., "Transcendental Multiproduct Production Functions," *International Economic Review,* vol. 5 (1964), pp. 273–284.
27. ———, and Razin, A., "On Multistage Multiproduct Production Functions, The Theory of Aggregation and Technical Change," unpublished manuscript, April 1967.
28. Nerlove, M., "Optimal Advertising under Dynamic Conditions," unpublished manuscript, July 1959.
29. Penrose, E. T., *The Theory of the Growth of the Firm,* Oxford, 1959.
30. Samuelson, P. A., *Foundations of Economic Analysis,* Cambridge, Mass., 1947.
31. Solow, R., "Technical Progress, Capital Formation and Economic Growth," *American Economic Review,* vol. 52 (1962), p. 76.
32. Stigler, G. J., "The Division of Labor is Limited by the Extent of the Market," *Journal of Political Economy,* vol. 59 (1951), pp. 185–193.
33. Tice, H. S., "Depreciation, Obsolescence, and the Measurement of the Aggregate Capital Stock of the United States, 1900–1962," *The Review of Income and Wealth,* June 1967, pp. 119–154.
34. Treadway, Arthur B., "Rational Entrepreneurial Behavior and the Dynamics of Investment," unpublished doctoral dissertation, University of Chicago, 1967.
35. Whitehead, A. N., *Science and the Modern World,* New York, 1925.

DISCUSSION

IRVING B. KRAVIS, University of Pennsylvania

Mr. Treadway uses his expanded definition of output to investigate certain problems in the theory of the firm. I would like in these remarks to consider the question of the definition of output in more general terms.

Many of the controversies about the nature of output and input center around two issues: first, the definition of the production boundary; and, second, the problems of measurement for outputs and inputs that change qualitatively over time.

Of these two sources of difficulty the old problem of finding an appropriate boundary of the economic process has been playing a lesser role. I can therefore avoid repeating, with lonely approval, the suggestions I made on this point at an earlier conference.[1] The issue keeps popping up nonetheless. Note Jean Wilburn's observation that production went down in barber shops as first the safety and then the electric razor transferred the daily production of smooth skinned male countenances from the barber shop to the home.[2] Current National Bureau studies of distribution have to cope with the fact that the spread of food supermarkets involves a shift of delivery, storage and other services formerly provided by the retailer to the provision of those services by the consumer himself. In both these cases, and many others like them, the official national accounts measures ignore the consumer's own activities in producing market analogues that contribute to the satisfaction of his wants. They assume also that man's capacity to transform goods into psychic satisfactions has been unchanging, or if not unchanging, immeasurable. All that we need to do therefore or all that it is possible to do—according to the prevailing view—is to record the change in the quantity of satisfaction yielding products that are delivered by the economy to man. If man becomes a more efficient consumer and requires a smaller quantity of a given good to achieve the same level of satisfaction as before that either cannot or should not be measured.[3] In Griliches' classic illustration, when dosage of pills necessary to prevent pregnancy was halved, output was and is measured, by the procedures in use, by the amount of the drug delivered to consumers and not in terms of the efficacy of each dose in achieving its purpose. After all, the resources absorbed per drug dose did not change and a halving of dosage was therefore equivalent to a halving of output, other things being equal.

1 Irving B. Kravis, "The Scope of Economic Activity in International Income Comparisons," *Problems in the International Comparison of Economic Accounts*, Studies in Income and Wealth 20, New York, NBER, 1957.

2 Victor R. Fuchs and Jean A. Wilburn, *Productivity Differences Within the Service Sector*, Occasional Paper No. 102, New York, NBER, 1967, p. 55 f.

3 For a comparison of the technology of consumption with the technology of production, see Kelvin Lancaster, "Change and Innovation in the Technology of Consumption," *American Economic Review*, May 1965.

This illustration brings us very quickly to the second and recently, more lively battlefront. Controversy here has sometimes reached such a fever pitch that charges of heresy have been heard. In the birth control case, it is possible to regard better knowledge as enhancing the consumers' efficiency as a pleasure machine, if it is not indelicate to use in this context a term employed recently by Fisher and Shell.[4] In most instances, however, the improvements flowing from technological progress fall more squarely in the production sector: television replaces radio; long-wear tires replace less durable ones; more efficient electric generators replace inferior varieties; etc. In these cases, the touchstone by which the change in output is measured is the change in input of resources.[5] If the new and more powerful generators take no more inputs to produce than the old ones, there has been no change in output.

I believe that on both these issues—the scope of economic activity and the treatment of quality change—the practice of national income accountants and of price index makers has become rigidified around compromises that were necessary and reasonable thirty years ago but can no longer be justified. What we should aim at doing has not changed; what we now can do has changed.

One way of seeing this is to ask the question, "If we could measure consumer satisfactions, would we wish to do it?" I think the answer to the question is "yes," and I believe that it follows that we should always have before us the objective of measuring the contribution of economic activity to human welfare. Of course, we cannot now measure satisfactions, but developments in sample surveys and in psychology have placed more within our reach the investigation and measurement of consumers preferences as between various goods. But even if we remain skeptical about rapid progress in this direction, we should keep the welfare goal before us. It will help us to avoid making an orthodoxy out of an early set of compromises, and to search continually for closer approximations to what we want.

[4] Franklin M. Fisher and Karl Shell, "Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index," in J. N. Wolfe (ed.), *Value, Capital, and Growth: Essays in Honour of Sir John Hicks,* Edinburgh, 1968.

[5] Milton Gilbert, "The Problem of Quality Changes and Index Numbers," *Economic Development and Cultural Change,* April 1961. Reprinted in *Monthly Labor Review,* September 1961.

One path to improvement, to which attention has been called in empirical work by Stone[6] and Griliches[7] and in theoretical work by Lancaster[8] as well as by some of the authors and discussants in this conference, is a more careful consideration of the choice of the physical units in terms of which individual products are priced and their output measured. This involves the recognition that the utility derived from a good is seldom unidimensional; most goods represent a cluster of characteristics that are sought by buyers. An orange is a combination of juiciness and sweetness; an automobile of comfort, power and economy; a knitting machine of speed, adaptability, labor requirements, repair costs and shutdown time. Many, though not all, quality changes over time (as well as quality differences at a given moment in time) consist of variations in the mix of these characteristics. When quality changes over time are significant for a product, it becomes important to catch these characteristics in our measurement. The more successful we are in selecting units that represent the characteristics really sought by buyers (and the less enmeshed we become in the intertemporal matching of purely physical specifications), the more successful we will be in catching the quality changes that are increases in output from a welfare standpoint but which elude present measures.

In the conventional method of price and quantity measurement the unit selected is that which is observed—one might say observed at first glance—in the market place: a car, a tractor, a physician's visit, a day in a hospital, etc. Differences in quality are either not taken into account at all or gauged on the basis of differences in the resources required to produce alternative qualities at a given moment in time. As Mr. Treadway's stress on disequilibrium conditions reminds us, this equivalence in cost does not necessarily mean an equivalence in consumer utility unless there is equilibrium in output and consumption as between the two product variants; if at the observed prices,

[6] Richard Stone, *Quantity and Price Indexes in National Accounts,* Paris, OEEC, 1956, p. 47 f.

[7] Zvi Griliches, "Notes on the Measurement of Price and Quality Changes," *Models of Income Determination,* Studies in Income and Wealth 28, New York, NBER, 1964. See also Irving B. Kravis and Robert E. Lipsey, "The Use of Regression Methods in International Price Comparisons," *International Economic Review,* forthcoming.

[8] Lancaster, "Change and Innovation in the Technology of Consumption."

the output of one is expanding relative to the other, that suggests that the former is preferred by consumers.

A more fundamental objection is that this conception of units does not probe deeply enough into market behavior. The presence on the market place of a variety of products that go under a given name such as "car" or "tractor" and sell for different prices indicates that the transactions between sellers and buyers must be specified in ways that are not taken into account by this gross approach. It is the business of measurement to identify these specifications that explain the variation in prices—whether they be gauged in pounds, horsepower, deadweight tons, kilowatts or combinations of these and others—and where possible to take these as the units of output. In this work the question that must be continually before us is "what qualities does the purchaser take into account in making his decisions among the alternatives before him?"

When these issues were discussed at an income conference five years ago, the multivariate approach to price and quantity measurement was described by one unenchanted observer as only "alternative statistical procedures for application of the conventional method." [9] I think that this statement is more true than not, provided the underlying conception of price and quantity measurement is the same when the two methods are used. The conventional method he had in mind was the ad hoc adjustment for quality, mainly in the price indexes, either on the basis of the relative prices of simultaneously available alternative qualities, or, when the various qualities are not on the market at the same time, on the basis of estimated differences in cost of production. However, the conventional approach is inferior to multivariate analysis because it is neither systematic nor comprehensive, and will often fall into arbitrary choices of criteria by which to adjust for quality change. In actual practice, furthermore, the choice of criteria by index number producers tends to turn upon differences in cost of production of physical features rather than on welfare-oriented indicators of output.

Even if such indicators were sought, a priori reasoning about the qualities that are important to the purchaser will usually lead to quali-

[9] Edward F. Denison, comment on Griliches' paper, in *Models of Income Determination,* p. 410. See also George Jaszi's comment, *ibid.,* p. 407 f.

ties at least some of which are difficult to quantify from a practical standpoint; in the case of a diesel engine, for example, the qualities might be power, economy of operation, reliability and durability. Perhaps in the longer run systematic ways to identify, measure and evaluate the relative importance of the relevant qualities should be sought. Meanwhile, multivariate analysis offers an approximating procedure; it consists of selecting in an objective manner from the physical characteristics in terms of which sellers and buyers describe the variants of the product, those that are important in explaining price variation. These are then taken as the characteristics—in the case of diesels, for example, they might be horsepower and weight—in terms of which changes in the output and price of the product are measured. The underlying assumption is that since the varying mix of these qualities explains the observed variation in price it can also be regarded as representing the qualities that purchasers are really seeking.

Although most of the empirical and theoretical work along these lines has been concerned with consumer goods, perhaps the most important possibilities for improved measurement of price and quantity changes lie in the capital goods sector. The multivariate approach would surely lead to a better resolution of the difficulties in this area with which writers such as Denison and Kendrick have wrestled in the past.[10] The long continuation of measurement of capital goods output mainly in terms of input can be ascribed to excessive reliance on the crutch "how much resources were absorbed?" and the failure to ask the more relevant question "what qualities does the purchaser seek?"[11] The biases in our measures of output in this critical area of GNP may be substantial. In the first place the stress on pricing physically identical goods from one period to the next greatly narrows

[10] Edward F. Denison, "Theoretical Aspects of Quality Change, Capital Consumption and Net Capital Formation," *Problems of Capital Formation: Concepts, Measurement, and Controlling Factors,* Studies in Income and Wealth 19, New York, NBER, 1957; and John W. Kendrick, "Some Theoretical Aspects of Capital Measurement," *American Economic Review,* May 1961.

[11] For a defense of input-based measures of the quantity of capital goods see Denison, *Problems of Capital Formation.* Denison rejects measures of the quantity of capital goods based on their output, but his argument rests on the use of a single output variable to measure the quantity of capital. What I am proposing is that we search for the variety of qualities that are important to the buyer; these might include, for example, directly or indirectly, the labor saving features of the new capital good.

the range of capital goods that can be included[12] and probably results in the systematic exclusion from the deflators of the kinds of goods most subject to technological change. Secondly, for the included goods quality adjustments tend to be made on a cost basis.[13]

The more fruitful approach which the availability of computers has made feasible is to regard each capital good as a cluster of characteristics being sought by purchasers (including such attributes as capacity or speed, requirements or saving of labor and other complementary inputs, durability, and repair time and costs) and to seek proxy variables for some or all of these qualities which are measurable and which account for the price variation observed in the market. Changes in the prices of these qualities, rather than in the prices of inputs or of some base period model which no longer exists, can then be estimated.

To enable us to select units of measurement from market data, buyers must be confronted with alternative specifications of a product, each with its own price. These conditions are frequently found in commodities but less often in services, for example, in education or government. There are, however, some service sectors to which the method could in principle be applied. In retail distribution, for example, gross margins of various kinds of establishment might be related to credit, delivery, and other aspects of customer service. Professor Reder's paper for this conference suggests that price indexes may be derived for medical care, since consumers are confronted with alternative packages of medical services, in the form of insurance plans, each with its own price tag. But where such objective bases for selecting units are not available, we cannot and do not now avoid subjective selection of units. For services as for commodities, one can conceive of a whole range of output measures ranging from those based on inputs such as hospital days to those representing utilities to consumers such as the prevention or cure of illnesses. The more closely the units of measurement represent the characteristics that the buyer

[12] In the U.S. it was reported a few years ago that the coverage of the wholesale price index was limited to 35 or 40 per cent of the commodities that enter into new investment in producers durables. Allan D. Searle, "Capital Goods Pricing; Problems and Prospects," Bureau of Labor Statistics, December 9, 1963, mimeo.

[13] *Idem.*

is really seeking, the more we will catch the output-increasing effects of quality change and the better will be our measures of output. Casting our measures of output in these terms would also resolve some of the issues about the production boundary in a different way; it would, for example, be less open to the claim that some improvements could be ascribed to the increased efficiency of the consumer (as has been suggested for the birth control pill case) and therefore beyond the realm of economic activity.

In our national accounts, thus far we have chosen input-based measures because they appear to be objective. We have in the past been warned against moving from these kinds of measures. "Switching the criterion for the 'commodity' to be priced from what the consumer actually buys (hospital care, surgeon's time, drugs, etc.) to what he really wants," it was said at an earlier conference, "is a dangerous and inconclusive game for the statistician to play." [14] But the conventional price or (output) measures do not completely escape decisions on the unit to be selected.[15] There are many products which can be and are quoted in alternative units, and the official measures must choose among them. Cotton sheeting, for example, may be reported in terms either of square yards or pounds (both are used in U.S. import statistics); the Bureau of Labor Statistics has chosen the per yard basis for its wholesale price index. Furthermore, danger and inconclusiveness have hardly been avoided by measuring prices and output in terms of units such as a "physician's visit" or a "hospital day"; it is no more justifiable in principle to take these radically changing services as units of output than it would be to take, for example, an "aluminum sheet" as the unit, regardless of changes in its quality and size. For both commodities and services, therefore, the alternatives are not, as the quotation seems to imply, objectively determined market units on the one hand and units chosen by individual introspection into what consumers "really want" on the other hand. In both areas, observation of the real nature of the transaction is important, and objective means of selecting the units should be pursued

[14] Denison, *Models of Income Determination,* p. 413. For a more recent defense of what might be called the first-glance basis for choosing the transaction unit, see Thomas W. Gavett, "Quality and a Pure Price Index," *Monthly Labor Review,* March 1967.

[15] This point was called to my attention by Robert Lipsey.

as far as possible. The final choice will, however, frequently be a matter of judgment, and the area of discretion will usually be higher in the case of services than commodities. The present choices of units largely evade the issue and represent one extreme of the range of subjective choices that can be made. Their inadequacy is shown by the fact that as soon as we begin to ask significant questions—as we have been doing in the medical care area—we are forced to get closer to welfare oriented measures of output such as Reder's quality adjusted index of the cost of medical care insurance or Mrs. Scitovsky's measures of the cost of specific illnesses.

I believe that it is necessary to reorient our price and quantity measures in a welfare direction. This will in fact lead to better measures of production, for there is no meaning to production unless it is considered in welfare terms. In any case, the most important policy uses of these measures are not, after all to measure production, but to tell us something about consumption and welfare. Perhaps the key area to which our price measures, for example, are relevant is policy with respect to inflation. The fear of inflation has been a great deterrent to a more expansionary monetary and fiscal policy in this country since World War II. But the good that price stability is held to do and the evil that it is held to avoid are largely matters of equity—that is, welfare. The appropriate price measures should be satisfaction oriented, as far as possible.

It will, of course, always be necessary to adopt conventions in economic statistics. They should, however, be continually reviewed for their relevance to the technical conditions of the day. I believe that technological change, particularly the computer, has outmoded some of our concepts about units of measurement and made excessive our reliance upon costs of production to resolve our difficulties. I do not claim that the approach I have advocated will solve all our problems; there are many types of quality change which will continue to elude us, to say nothing of our failure to measure the welfare effects of increased variety.[16] I do claim, however, that I have pointed to the right question and that keeping it before us can lead to great im-

[16] See J. L. Nicholson, "The Measurement of Quality Changes," *Economic Journal,* September 1967.

provements in our present practices, particularly with respect to capital goods and consumers durables but not excluding the services.

TA-CHUNG LIU, Cornell University

In his valuable contribution, Treadway first reviews the familiar weaknesses of the conventional methods of measuring output; the inadequate representation of output by value added; and the unsatisfactory measurement of factor productivity by income shares on the assumptions of perfect competition, existence of equilibrium, constant returns to scale, and the absence of externalities.

The positive contribution of the author lies in his very ingenious and formal development of a dynamic theory of the firm, giving explicit and detailed treatment to disequilibrium and variable returns to scale.

A fairly general dynamic model of the firm is sketched in Section II; a full treatment of the model is given to a specialized version of the model with the following assumptions: one output (Q, with price P), one perfectly variable input (L, with price W), one imperfectly variable factor (K, with "rental" R and external price G), and an interest rate (r).

The firm is assumed to maximize the following present value:

$$\int_{\substack{0\\L,K}}^{\infty} [PQ - WL - RK - G\dot{K}]e^{-rt}\, dt$$

subject to the production function:

$$Q + C(\dot{K}) - F(K, L) = 0$$

and the initial condition $K(0) = K_0$.

Equivalently, the firm attempts to maximize:

$$\int_{\substack{0\\L,K}}^{\infty} [PF(K, L) - WL - RK - G\dot{K} - PC(\dot{K})]e^{-rt}\, dt$$

NOTE: These comments are based on the original version of Treadway's paper as presented at the Conference.

The two novel features of this model are: 1. The inclusion in the production function of two outputs: the good it sells (Q) and internally accumulated capital ($\dot{K}$). 2. The allowance of variable returns to scale. Both features are indeed desirable and interesting.

Since internal accumulation ($\dot{K}$) contributes to the net worth of the firm, one wonders why the firm would want to subtract the terms involving $\dot{K}$ from the integrand. The answer lies in the assumption that the horizon of the firm extends to the infinite future. A more limited horizon with a terminal preference would be more realistic and interesting. The techniques of analysis used here (the calculus of variations) is adopted from the pathbreaking paper by F. P. Ramsey. (A readily accessible summary of Ramsey's paper, together with a clear explanation of the technique of calculus of variations, can be found in G. D. Allen, *Mathematical Analysis for Economists,* Chapter XX.) In Ramsey's paper, a terminal condition is assumed. It would be of interest to see what effects different assumptions about the terminal condition would have on the result. For instance, a firm may be interested in the maximization of profits subject to the production function and a certain rate of growth of capital assets during a finite period of time. This seems to be a rather popular attitude of the business world in the postwar period. The terminal conditions would at least affect the arbitrary constants that appear on integration. It would be interesting to explore into the economic implications of such conditions.

Finally, it is not obvious how in fact the elaborate theoretical model developed by the author would help in the measurement of output which is presumably the main task of this paper. The author mentioned that he has developed "an analytical framework for which parametric representations suitable for empirical work can be designed." (Page 78.) What are the parameters? How to represent them in empirical work? I think the author is on the right track, but one would like to seek more explicit guidance for measurement purposes.