EG & ET

WORKING PAPERS

PROCESS ANALYSIS, CAPITAL UTILIZATION, AND THE EXISTENCE OF DUAL COST AND PRODUCTION FUNCTIONS*

Christopher C. Klein

WORKING PAPER NO. 116

May 1984

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

PROCESS ANALYSIS, CAPITAL UTILIZATION, AND THE EXISTENCE OF DUAL COST AND PRODUCTION FUNCTIONS*

Christopher C. Klein

Federal Trade Commission

Bureau of Economics

Washington, D.C. 20580

March 1984

* Thanks are due to Knox Lovell, John Stewart, Helen Tauchen, and, especially, Nicholas Georgescu-Roegen for their invaluable comments. The views expressed here are those of the author alone and do not represent those of the Federal Trade Commission.

I. Introduction

This paper integrates the literature on capital utilization and shiftwork with the duality theory of cost and production functions. The model constructed below posits a production technology consisting of an instantaneous rate of production function and a time duration variable. The existence of a cost function dual to this technology then follows directly from assumptions on the mathematical properties of the rate function.

Marris's (1964) original work on the economics of capital utilization, and the more recent contributions of Betancourt and Clague (1981), and Winston (1982) either ignore or are outright hostile toward duality theory. The time utilization of plant and equipment embodied in these models, however, is not inconsistent with the mathematical methods used in the "timeless" duality theory. Indeed, the model proposed here contains neoclassical duality theory as a special case.

The crucial tool in this construction is Georgescu-Roegen's (1970, 1971, 1972) process analysis of production. His analysis of an idealized factory yields a production model that precisely specifies the inputs and initially recognizes the time duration of production. Thus, the production process is directly modeled, rather than abstractly constructed. In this way the model is similar to engineering approaches to production.[1]

The interesting aspects of the model, however, derive not from the process analysis, but from the specification of input prices. The time element enters the cost function because the input prices must account for the time use of the inputs. Time affects the cost of labor directly, while the purchase price of capital is generally invariant with respect to production time. This asymmetry in the behavior of input prices over time combined with the returns to scale characteristics of the rate function determine the cost minimizing length of the "working day".

For the sake of argument, the term "neoclassical" is used to denote the class of production models which ignore the firm's choice of the working day. These models treat inputs as flows accumulated over a fixed period of time with fixed (time invariant) prices. Recent theory has employed these models to such an extent that Winston criticizes them under the heading of duality models. Nevertheless, the objectionable assumptions underlying this literature can be traced to Wicksteed's pioneering work in the Nineteenth Century.[2] Thus, neoclassical is used here, however unfairly, to characterize a diverse literature in production theory.[3]

Fortunately, the mathematics behind duality approaches to production are blissfully indifferent to the meanings we attach to them.[4] This allows us to apply familiar mathematical concepts to a novel production technology with relative ease. In fact, it is possible to derive results obtained by Winston, and by Betancourt and Clague for linearly homogeneous production functions and extend them to the class of homothetic production functions.

For purposes of exposition, the model is built in reference to a working day. This is not a prerequisite. The time period to be determined can be structured in any number of ways. Daily production, however, is immediately applicable to questions of capital utilization and shiftwork on which the existing literature has focused.

The process analysis of factory production is reviewed in the following section. Section III establishes the cost function dual to the production technology of section II, and demonstrates some graphical properties of the model using familiar isoquant concepts. Section IV extends the model to include time variable prices and a sloped output demand curve in the context of profit maximization. Section V examines the characteristics of the optimum length of the working day and section VI concludes the investigation.

II. Factory Production: A Process Analysis View

Suppose we wish to analyze the production process of an industrial factory. By the term factory is meant an assembly line or "in line" process as opposed to an "in parallel" process such as agriculture. In an assembly line factory there are a succession of work stations such that at any given time there is a potential unit of output in each work station. Furthermore, each worker and each article of capital are continuously employed by switching to a new unit of output as soon as their task on the previous unit has been completed. What we observe is a continuous line of "goods in process" moving successively from work station to work station until exiting the process as finished goods.

The process analysis of a factory begins by surrounding the factory with an analytical boundary. This boundary is an imaginary barrier placed around the factory at the discretion of the investigator in order to separate the process to be observed from the rest of the world. The investigator may also choose the placement of this boundary as it best suits the goals of the analysis. Let us choose our analytical boundary such that all raw materials, energy, and previously processed (i.e., intermediate) inputs must move across this boundary directly into the factory process, and all outputs cross the boundary immediately upon exit from the process. This has been done schematically in Figure 1.

We can observe that at any moment there is some flow of inputs across the boundary and some associated flow of output and "waste products" across the boundary in the opposite direction. We also observe some factors of production present inside the boundary whenever the factory is in operation. These factors are what we commonly think of as labor and capital, plus the goods in process present at each work station along the assembly line. Thus the flow of output produced is related to the flow of material and energy inputs and the "quantity" of labor and capital present at the time. In mathematical notation

$$(1) \qquad q(\tau) = f(K(\tau),\ H(\tau),\ x(\tau))$$

says that the maximum rate at which output crosses the process boundary at time τ, $q(\tau)$, is a function of the capital K and labor H present at τ, and the rate x at which other inputs enter the

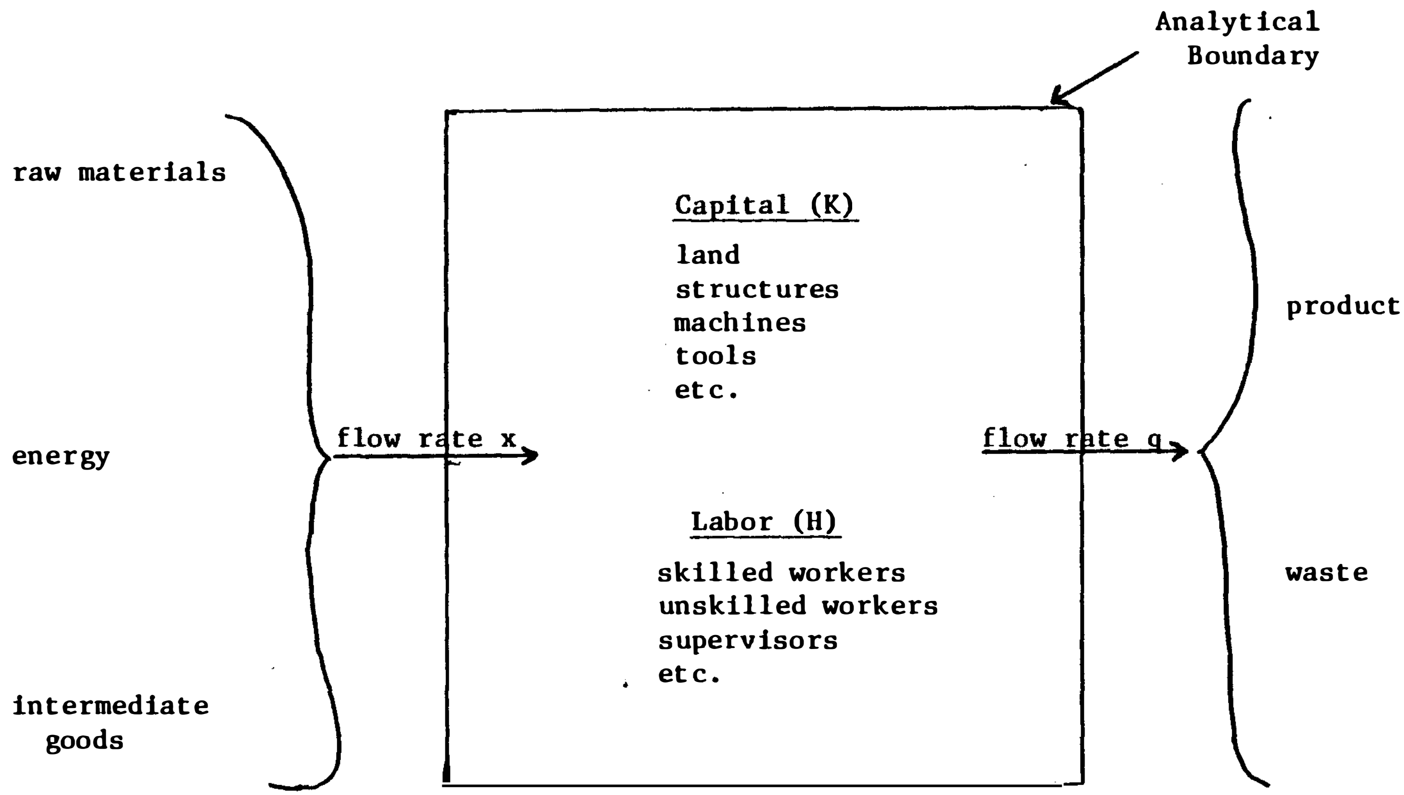


Fig. 1. Process Analysis of a Factory

process at time τ. Notice that q can be a vector of output rates, excluding the residual waste products, for the case of joint saleable outputs; that K is a census of machines and structures of each types present at τ; that H is a census of workers of each type present at τ; and x is a vector of input flow rates at time τ.

If T is the period over which the factory operates, then total production is given by

$$(2) \qquad Q = \int_0^T f(K(\tau), H(\tau), x(\tau))d\tau.$$

If the capital and labor present over the period and the flow rates of inputs are constant, and the productivities of the factors are not affected by the length of the period, we can simplify (2) to

$$(3) \qquad Q = T \cdot f(K, H, x).$$

This says that the total output of the factory is determined by the length of time the factory operates T, the capital and labor on hand over the period, and the flow of other inputs over the period. In addition (3) indicates that in order to produce the quantity Q of product the firm must choose the capital and labor to hire, the flow of inputs to purchase, and the length of the period to operate, T. Therefore the process analysis of factory production provides us with a production model that requires the firm to choose input factor proportions as well as a time period of operation in order to produce the desired output quantity Q.

Equation (3) also shows an unfamiliar relation between the input factors and the quantity of output. K and H are just numbers of men and machines independent of time, while x is a vector of flow rates, or quantities per unit time. The first to develop a model accounting for, and elaborating on, the differences among capital, labor, and flow inputs was Nicholas Georgescu-Roegen (1971). He denoted the capital and labor components of (3) by the term <u>fund factors</u> and the material and energy inputs by the term <u>flow factors</u>. The derivation of the flow factor's nomenclature is fairly obvious from (3) since the flow factors appear as flows in the function f. The case for the fund factors is not, in comparison, clear.

Georgescu-Roegen reasons that the stock of capital inputs is actually a stock of machines that is perpetually maintained by the process in which it participates. This stock represents the available capital services the firm buys when it purchases a machine, or a whole plant, where the actual amount of capital services used is measured as machine hours for each type of machine. This is less misleading than talking of service stocks because there is no physical flow that augments or depletes the capital fund. In fact, the capital fund can only be drawn down by running the machine until it is no longer economically feasible to do so. One cannot consume the entire fund immediately nor stretch the fund over a longer period by feeding in new supplies of services. Regular maintenance must be undertaken to insure each machine's operating efficiency, of course, but this does not increase the fund represented by a machine. At some point the machine will become very expensive to maintain, at which time it will be scrapped.

There are two major characteristics of a fund factor of production that we can identify:

1. The fund factors are not quantitatively altered by the production process.
2. A period of time is required to exhaust the services represented by a fund factor.

It is obvious that the capital and labor elements of (3) represent fund factors in the factory process. Moreover, it is the <u>rate of fund service</u> that appears in (3) and is denoted by K and H. This can be illustrated with a simple example. Suppose a factory employs 50 workers for a period of 10 hours. The quantity of labor services used over the period is given by (50 men)·(10 hours) = 500 manhours. The rate of service use is then the quantity of labor services used divided by the period of time over which the services are rendered, or (500 manhours)/(10 hours) = 50 men. The rate of fund service is just the number of each fund factor present at any time the factory is in operation.

Now reconsider the right hand side of (3). Since K and H are the census figures for machines and workers present while the factory operates, they are the number of each type of fund factor present, therefore they represent the rates of fund service. In addition, the vector x represents the observed rates of flow of material and energy inputs into the process. This indicates that the function f(K, H, x) relates the rate at which output is produced to the rates at which input factors are used (as in (1)).

It may be instructive to consider two comments on factory processes. The first is by Georgescu-Roegen as he explains the fund factors participation in the process, ". . . in the case of manufactured or mined products, we can arrange the elementary processes in line in such a manner that each fund shifts to another process as soon as it has finished its task in the previous one. This is how any factory operates, like an assembly line even though one is not in direct view."[5] A similar point has been mentioned by Marsden, Pingry, and Whinston, ". . . an assembly plant could be characterized by a series of assembly units (reactors), and the process could be described in terms of a rate equation."[6] This is precisely what equation (1) and its counterpart equation (3) attempt for a factory process. The rate of output is determined by a function of input rates, and the total quantity is found by multiplying the rate of output by the time period T.

Georgescu-Roegen spent a great deal of time formulating the substitution relationships among the inputs so that their compliance with the physicists' principle of Conservation of Matter-Energy was clear. This expanded the number and complexity of the structural equations and emphasized the unusual aspects of the flow-fund approach. While his treatment of the waste outputs and the relation of the inputs and outputs to plant capacity was uniquely insightful, it undermined the intuitive appeal and the empirical utility of his production model.

Fortunately it is possible to construct a model in the spirit of Georgescu-Roegen that can be represented more simply. Equation (3) and a dose of Alfred Marshall's "sensitiveness of touch" are the main ingredients. The primary difficulty is in understanding the way in which substitution occurs between the flow and fund factors and its relation to the conservation principle. Examine equation (2) once again

$$q = f(K, H, x).$$

By the conservation principle we know that in matter-energy terms the flow inputs must equal the flow outputs. Suppose we define q to be only the product flow, or a restricted output vector that excludes the waste products. If we are given (2) and some values K, H, and x, we can find q. We can then calculate the "waste" output from q and x: $w = [x-q]$. Since we can never observe a process that defeats the conservation principle, it is not necessary to carry it as an explicit constraint. We can presume that any process we might meet obeys all physical laws and subsume these implicitly in the f function.

There are several advantages to this approach. First, this allows the rate function f to display all the usual substitutability properties we are accustomed to, except in terms of rates instead of quantities as in the strict neoclassical world. If we wish to increase the rate of output of the product by hiring more labor but no more material inputs, then this requires that the additional labor alter the relative amount of waste product flows. If the waste flow is not altered then the rate of product flow will not change, but the f function still gives us a maximum output rate q for any values of K, H, and x.

Furthermore, we can define the "product" and the "waste" outputs by letting q be the outputs with positive prices. The waste products are then identified as the residuals.

In summary, we have a production model in (2) and (3) that gives the rate of output as a function of rates of input flows and rates of fund service. This model allows substitution among the inputs similar to the usual neoclassical production function, while incorporating the strengths of the flow-fund model.

III. Duality In the Flow-Fund Production Model

Suppose we have an N factor neoclassical production function F: $u = F(y_1, y_2, \ldots, y_N) = F(y)$ where u is the amount of output produced during a given period of time and $y = (y_1, \ldots, y_N) > 0_N$ is a nonnegative vector of input quantities used during the period. Suppose also that the producer faces fixed positive prices for inputs $(p_1, p_2, \ldots, p_N) \equiv p$ and that the producer does not possess market power in the input markets.

We now define the producer's *cost function* C as the solution to the problem of minimizing the cost of producing at least output level u, given the input prices p, or

(4) $$C(u, p) = \min_y \{p'y : F(y) \geq u\} .$$

Assumption 1 below is sufficient to imply the existence of solutions to the cost minimization problem as stated in the lemma that follows.

Assumption 1: F is continuous from above; i.e., for every $u \in$ range F, $L(u) \equiv \{y : y \geq 0_N, F(y) \geq u\}$ is a closed set.

Lemma 1: If F satisfies Assumption 1 with $p \gg 0_N$, then for every $u \in$ range F, $\min_y\{p'y : F(y) \geq u\}$ exists.

Furthermore the following seven properties can be derived for the cost function C requiring only Assumption 1 on F.

Properties I for C:

C1: C is a non-negative function; i.e., for every $u \in$ range F and $p \gg 0_N$, $C(u, p) \geq 0$.

C2: C is (positively) *linearly homogeneous* in input prices for any fixed level of output; i.e., for every $u \in$ range F, if $p \gg 0_N$ and $k > 0$, then $C(u, kp) = kC(u, p)$.

C3: If any combination of input prices increases, then the minimum cost of producing any feasible output level u will not decrease; i.e., if $u \in$ range F and $p_1 \geq p_0$, then $C(u, p_1) \geq C(u, p_0)$.

C4: For every $u \in$ range F, $C(u, p)$ is a <u>concave</u> function of p.

C5: For $u \in$ range F, $C(u, p)$ is <u>continuous</u> in p, $p >> 0_N$.

C6: $C(u, p)$ is <u>non-decreasing</u> in u for fixed p; i.e., if $p >> 0_N$, u_0, $u_1 \in$ range F and $u_0 < u_1$, then $C(u_0, p) < C(u_1, p)$.

C7: For every $p >> 0_N$, $C(u, p)$ is continuous from below in u; i.e. if $p^* >> 0_N$, $u^* \in$ range F, $u_n \in$ range F for $n = 1, 2, \ldots$, $u_1 < u_2 < \ldots$ and $\lim u_n = u^*$, then $\lim_n C(u_n, p) = C(u^*, p^*)$.

These results are well known and the proofs can be found in numerous places. We have followed Diewert's (1982) rendition of the standard duality results above and follow his work once again in stating the next result, frequently called Shephard's Lemma.

<u>Lemma 2</u>: If the cost function satisfies Properties I for C and, in addition, is differentiable with respect to input prices at the point (u^*, p^*), then

$$(5) \qquad y(u^*, p^*) = \nabla p\, C(u^*, p^*)$$

where $y(u^*, p^*) \equiv [y_1(u^*, p^*), \ldots, y_N(u^*, p^*)]'$ is the vector of cost minimizing input quantities needed to produce u^* units of output given input prices p^*, where the underlying production function F is defined above, $u^* \in$ range F and $p^* >> 0_N$.

The advantage of using Lemma 2 is that only the cost function is required to have certain properties and the corresponding production function F is derived from the given cost function. Given a cost function satisfying Properties I, we can derive the input demand equations directly from the functional form of C. It is not necessary to find F nor to carry through the entire constrained maximization of the production function explicitly to find the input demand equations.

Now suppose we have the production rate function below:

(6) $q = f(K, H, x)$.

where q is the maximum quantity produced by continuous operation of a factory at a constant instantaneous rate for a 24 hour period; K is the rate of capital fund service over the period, or the number of machines and structures present over the period, assumed constant for any "day"; H is the rate of fund service for labor, or the number or workers of each type present at any time during the period; and x is the quantity of material inputs used for a constant instantaneous rate of input flow sustained for a 24 hour period.

Notice that equation (6) describes a production technology commensurate with the neoclassical production function when the given period of time is 24 hours and the measurement of labor and capital inputs is specified more precisely. However, the quantity produced in a working day is given by

(7) $Q = t \cdot f(K, H, x), \quad 0 < t \leq 1$

where t is the utilization rate, or proportion of a 24 hour period that the factory is in operation, and it follows that $Q = tq$.

Since a dual cost function must exist for f in (6) with given input prices by Lemma 1, a cost function dual to (7) should also exist, but it will not be identical to the neoclassical cost function due to the institutional differences in payments to capital vs. the other inputs. We take account of the firm's ability to purchase labor services and material input flows, but only the capital fund in the definition of the flow-fund cost function

$$(8) \qquad C°(Q, t, p) \equiv \min_{K,H,x} \{p_k K + tp_h H + tp_x x : tf(K,H, x) \geq Q\}$$

or equivalently

$$(9) \qquad C(q, P) \equiv \min_{K,H,x} \{p_k K + tp_h H + tp_x x : f(K, H, x) \geq q, \; tq = Q\}$$

where P is the modified input price vector, $P \equiv (p_k, tp_h, tp_x)$.[7] Since for any given value of t, $tf(K, H, x) \geq Q$ implies $f(K, H, x) \geq q$ where $Q = tq$ by definition, the two forms of cost (8) and (9) are identical for any quantity Q, and given values for t and input prices p.

The equivalence of (8) and (9) causes some difficulties in notation when the derivatives of C and C° with respect to q and t are encountered. Since this occurs repeately, the following conventions in notation are followed:

$$(10) \qquad \begin{aligned} C_q &= \partial C/\partial q = dC°/dq \Big|_{\substack{p=p_0 \\ t=t_0}} = \partial C°/\partial Q \; (\partial Q/\partial q) = (\partial C/\partial Q)t \\ C_t &= \partial C/\partial t = dC°/dt \Big|_{q=q_0} = \partial C°/\partial Q \; (\partial Q/\partial t) + \partial C°/\partial t \end{aligned}$$

The usual notation for a partial derivative is used throughout. Note that prices are not required to remain fixed for C_t since we will presently allow some prices to change with utilization rate. We interpret C_q and C_t as the marginal cost of increasing output by increasing rate of production and the marginal cost of increasing output by increasing utilization rate, respectively.

Recall that the input prices in (8) - (10) are daily prices, or prices per 24 hours. This means that p_k is depreciation and interest for one day; p_h is the cost of employing one worker for 24 hours, or $p_h = 24w_h$ where w_h is the hourly wage; and p_x is just the price per unit of the flow inputs.

Suppose that C and C° have properties I for C, so that factor demand equations can be found by Lemma 2. Suppose further that C(q, P) is twice continuously differentiable at the point (q*, P*). Then differentiating (9) with respect to factor prices gives daily factor demand functions

$$\partial C/\partial p_k = K(q^*, P^*), \quad \partial C/\partial p_h = tH(q^*, P^*),$$
$$(11)$$
$$\partial C/\partial p_x = tx(q^*, P^*).$$

Further, we know the following so-called symmetry conditions must hold[8]

$$(12) \qquad \partial K/\partial p_h = t(\partial H/\partial p_k), \; t(\partial H/\partial p_x) = t(\partial x/\partial p_x), \; \ldots.$$

But here is a surprise. For the labor and flow inputs we get the usual result

(13) $\partial H/\partial p_x = \partial x/\partial p_h$,

while for the capital input we find

(14) $\partial K/\partial p_h = t(\partial H/\partial p_k)$, $\partial K/\partial p_x = t(\partial x/\partial p_k)$;

i.e., the usual conditions are now weighted by the utilization rate t. Georgescu-Roegen (1972) derives an equivalent result using the Lagrangean method: that the marginal rate of substitution of capital for labor is equal to the price ratio weighted by the utilization rate. He interprets this to mean that the budget line in K and H space is not tangent to the relevant isoquant when $t < 1$. We know that this is not possible for a (K, H) combination that is a solution to the cost minimization problem. A closer look at (14) resolves this paradox.

Recall, first of all, that p_h is defined as a daily price for labor, $p_h = 24w_h$ when w_h is the hourly wage. This means changes in p_h arise from changes in w_h alone, or

(15) $dp_h = (24)\cdot dw_h$.

And using this with (14) yields

(16) $(\partial K/\partial w_h)(1/24) = t(\partial H/\partial p_k)$,

which we rearrange to find

(17) $\partial K/\partial w_h = 24t(\partial H/\partial p_k) = T(\partial H/\partial p_k) = \partial(TH)/\partial p_k$,

where T is the number of hours per day that the factory operates. Of course TH is nothing more than the number of manhours, or the quantity of labor services, used by the firm in a working day. We can now write the demand for labor services as

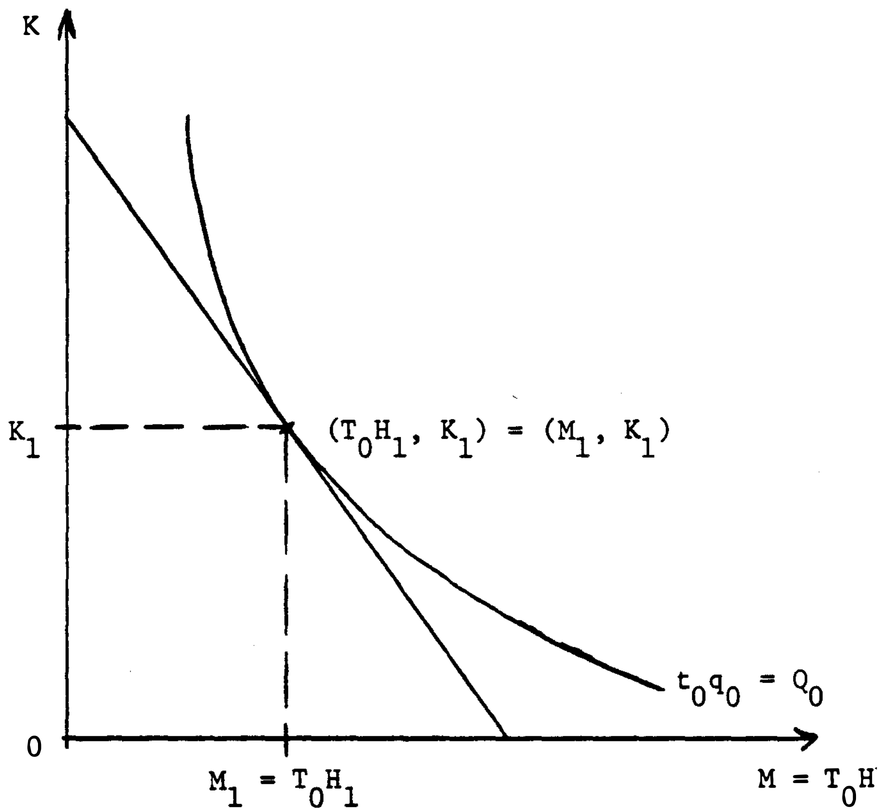
(18) $M(T, q, P) = T \cdot H(q, P)$,

and this allows us to rewrite (14) in the more familar form

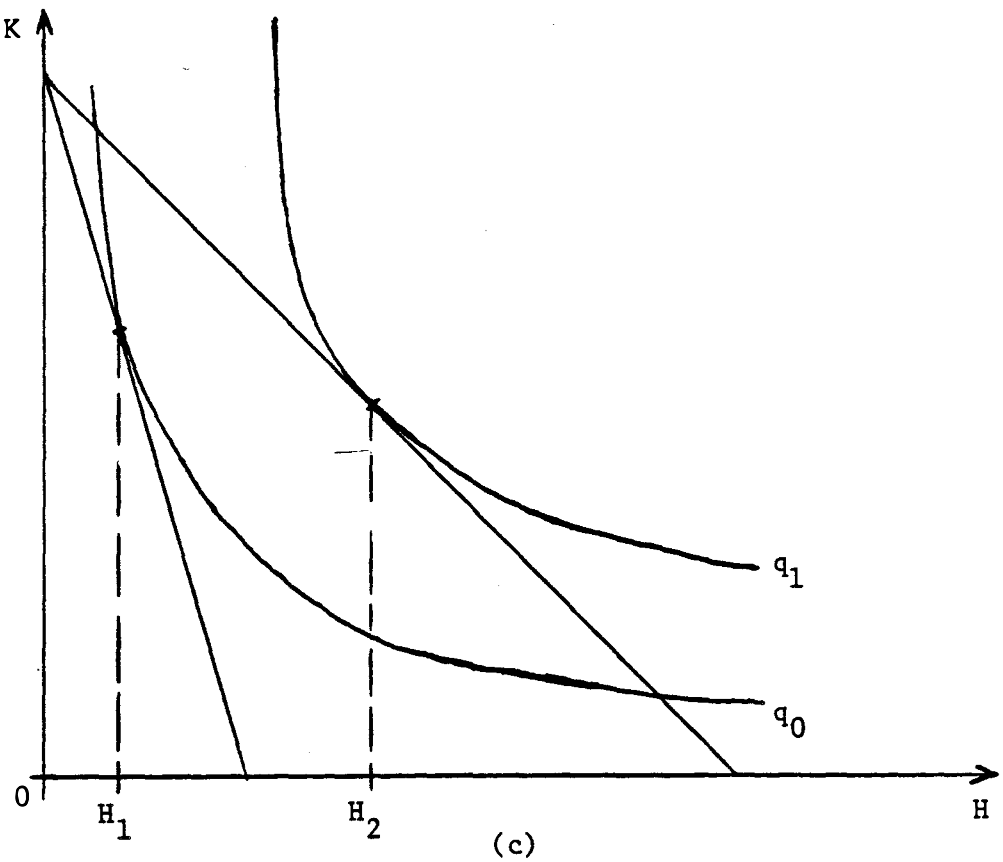
(19) $\partial K/\partial w_h = \partial M/\partial p_k$.

Clearly the relevant isoquant is in K and M = TH space as shown in Fig. 2 (a). Furthermore, the isoquants contain both a production rate and a utilization rate component, $Q = t \cdot q$. In 2(a) the isoquant is drawn for a fixed utilization rate t_0, and the labor services determined by the tangency imply some number of workers, H_1, for the entire working day, $T_0 = 24t_0$, so $M_1 = T_0H_1$.

If we wish to consider utilization rates other than t_0, it is necessary to redraw the isoquant map for the new t. This has been done in Fig. 2(b) for $t_1 < t_0$. Reducing the length of the working day from T_0 to T_1 shifts the isoquants so that the firm moves to a higher rate q_1 to maintain output at Q_0. Figure 2(c) shows this change in the utilization rate in K and H space. Here the fall in the length of the working day appears as a fall in the "price," or relative daily cost, of a worker which rotates the budget line away from the H axis. The firm reacts to this change as it would to a change in the wage, by substituting labor for



(a)



(c)

Figure 2

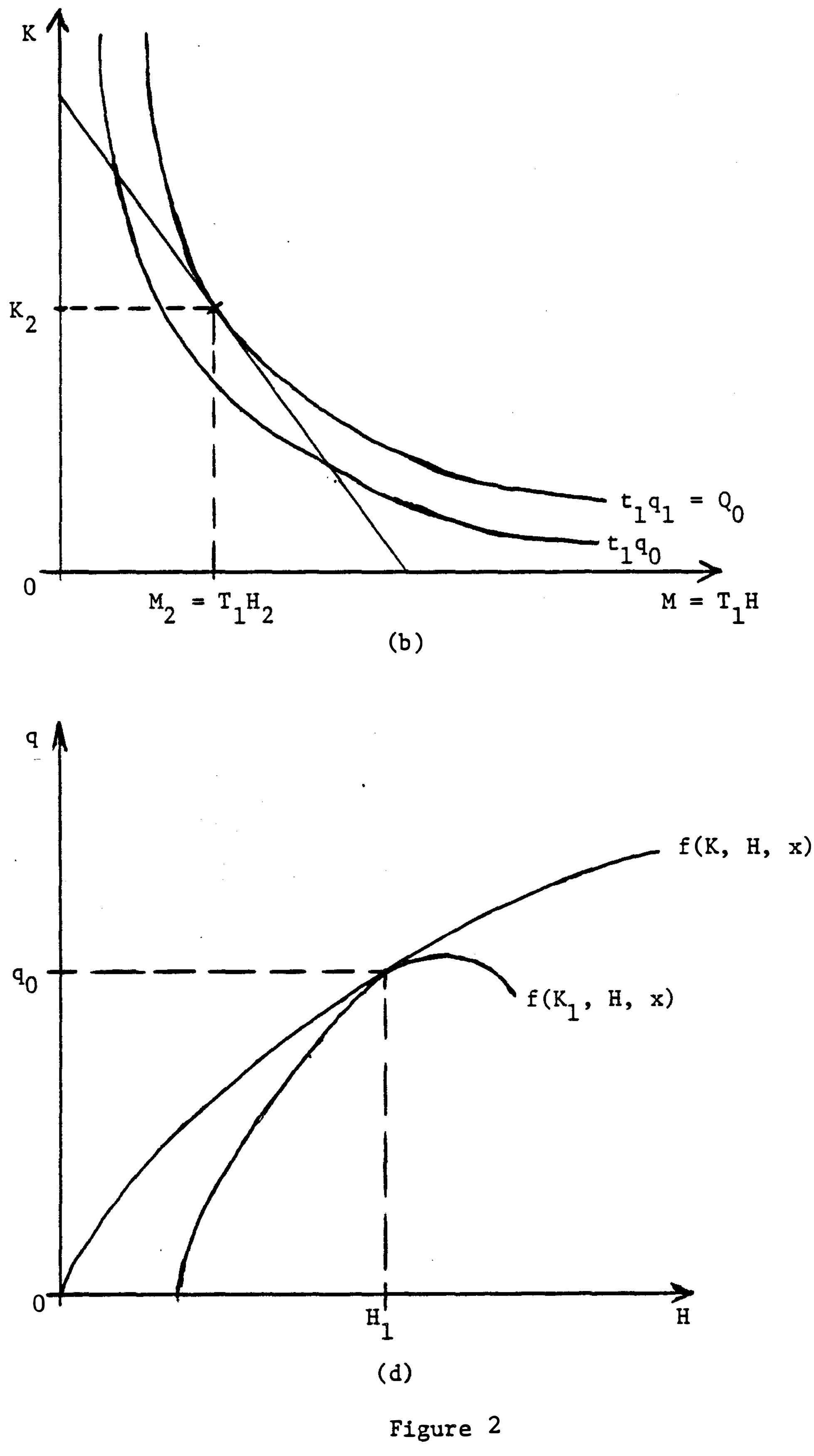


Figure 2

capital and moving to a higher rate of production. We can see that the deciding factor for the firm is no longer the relative prices of the inputs, but the relative daily cost of the inputs when the length of the working day is a decision variable.

Suppose the producer chooses production rate q_0 and hires K_1 and H_1. This represents only one point on the plant's planning, or long run, production rate function $f(K, H, x)$. The producer is free to vary his labor and flow inputs on a daily basis so he has the short run production rate function $f(K_1, H, x)$ shown in Fig. 2(d). The firm is always able to move along $f(K, H, x)$ at any time, but can forego the cost of rearranging its capital from day to day by moving along $f(K_1, H, x)$ as demand fluctuates. Now we may observe the firm adjusting its input mix as well as the length of the working day in response to short run fluctuations.

IV. Variable Output and Input Prices

Two obvious extensions of the model involve the usual "monopoly problem", in which the firm faces a sloped demand curve, and the case of input prices that vary with the utilization rate. The derivations of both of these are based on Diewert (1982), where the case of a variable input price is formally a modification of the monoposony problem. This section derives the flow-fund equivalents of the neoclassical duality results.

Suppose we have a flow-fund production function f that satisfies the following conditions:

(i) f is a real valued function defined over the non-negative orthant and <u>continuous</u> on this domain.

(ii) f is <u>increasing</u>.

(iii) f is a <u>quasiconcave</u> function.

Then given f and positive inputs and outputs, we write the profit maximization problem as

$$(20) \qquad \max_{t,q} \{D(Q)Q - C(q, P) : Q = tq > 0,\ 0 < t \leq 1\}$$

$$= D(Q^*)Q^* - C(q^*, P^*).$$

Q is total daily output, D(Q) is a daily inverse demand function, and C(q, P,) is the cost function dual to f. Solutions to the profit maximization problem are denoted by an asterisk. The first order conditions for a maximum are

$$(21) \qquad \begin{aligned} D(Q)q + (\partial D/\partial Q)qQ - C_t &= 0 \\ D(Q)t + (\partial D/\partial Q)tQ - C_q &= 0 \end{aligned}$$

where C_q and C_t are defined by (10). This condition can be restated as

$$(22) \qquad C_q/t = MR = C_t/q$$

where MR is marginal revenue in the usual sense. This condition can be interpreted as a variant of the typical neoclassical marginal condition: the firm chooses to operate where the marginal gain from an increment to time in production is just equal to the marginal gain from an increment to rate of production. C_q/t and

C_t/q are "marginal cost" in the two output dimensions weighted by the reciprocals of t and q. Thus "marginal revenue equals marginal cost" in (30).

Input prices that vary by "time of day" can be incorporated in this framework, if the utilization rate can be measured relative to a specific starting time during the day. The proper measurement of t will guarantee that the size of t corresponds to a known set of operating hours such that t and "time of day" are paired. In this context input prices that change over the course of the day can be described by a price function P, $P(t) > 0$, that depends on utilization rate.

Applying Diewert's (1982) treatment of the monopoly and monopsony problems for econometric purposes, if we know the output demand function and the input price or supply functions, then the system defined by (21) is identical to the profit function dual to f when the output and input prices have been "linearized" as their shadow prices. That is, for example, if

$$p_q = D(Q) + [\partial D(Q)/\partial Q]Q = P^*_Q + D'(Q)Q,$$

is the shadow or marginal price of output, where P^*_Q is the product price, then

$$\max_{q,t} \{p_q Q - P'X \; ; \; Q = tq > 0, \; 0 < t \leq 1\} = \Pi(p_q, P)$$

where Π is the firm's true profit function which is dual to the rate function f.[9]

V. Restrictions on the Production and Cost Functions

We can now make some statements on the firm's optimal choice of q and t by imposing restrictions on the form of f, which imply a form for C. Suppose f is homogeneous of degree $1/\theta$ such that

$$(23) \qquad f(X) = [g(X)]^{1/\theta}$$

where g(X) is linearly homogeneous. Then

$$f(\lambda X) = \lambda^{1/\theta} f(X) = [g(\lambda X)]^{1/\theta} = [\lambda g(X)]^{1/\theta} = \lambda^{1/\theta} [g(X)]^{1/\theta}$$

and the cost function dual to f decomposes in the following way.

$$(24) \qquad C(q,P) = q^{\theta} C(1,P) = q^{\theta} c(P)$$

where $\theta = 1$ implies constant returns to scale, $\theta < 1$ implies increasing returns to scale, $\theta > 1$ implies decreasing returns to scale, and c(P) is the unit cost function dual to g.

If we impose the neoclassical assumptions of constant returns to scale and perfectly competitive firms, the first order conditions for profit maximization become

$$(24) \qquad \frac{c(P)}{t} \qquad P_Q^* = \frac{\partial c(P)}{\partial t}.$$

If all factors of production are paid only for the time they actually participate in the process, t, and all factor prices are constant over all values of t, then P = tp and

$$c(P) = c(tp) = tc(p)$$

by the homogeneity property of c in P. The first order conditions then become

$$(26) \qquad c(p) = P^*_Q = c(p) \ .$$

In this case any combination of t and q will do, since all combinations have identical marginal cost. This occurs because we have ignored the differences in factor payments that distinguish the flow-fund model from neoclassical production models. (Note that we can derive (26) by assuming $t = 1$.) Thus the neoclassical ignorance of the time factor in production, t, is perfectly consistent with ignoring the differences in factor payments. In fact, $q \cdot c(p)$ is the neoclassical cost function dual to f.

The conditions required for the firm to choose $t < 1$ can now be derived from the firm's average daily cost function,

$$(27) \qquad A(q,P) \equiv C(q,P)/tq = q^{\theta-1}\, t^{-1} c(P) \quad .$$

Totally differentiating and rearranging (27) gives

$$(28) \qquad \left.\frac{dA}{dt}\right|_{dQ=0} = \frac{\partial A}{\partial q}\frac{dq}{dt} + \frac{\partial A}{\partial t} = \frac{\partial A}{\partial q}\frac{-q}{t} + \frac{\partial A}{\partial t} \quad ,$$

where we have used $dQ = tdq + qdt = 0$. Equation (28) shows two effects. The first is the change in the rate of production required to hold the total quantity, Q, constant as t changes. The second is the effect of changing input prices on factor proportions and average cost caused by changes in t. Carrying out the differentiation yields

$$\left.\frac{dA}{dt}\right|_{dQ=0} = [(\theta-1)q^{\theta-2}\ t^{-1}c(P)](-q/t) + q^{\theta-1}\ t^{-1}\partial c(P)/\partial t$$
$$- q^{\theta-1}\ t^{-2}c(P)$$

which reduces to

(29) $$\left.\frac{dA}{dt}\right|_{dQ=0} = (q^{\theta-1}\ t^{-1})\ [\partial c/\partial t - \theta c(P)/t]$$

and produces the condition

(30) $$\left.\frac{dA}{dt}\right|_{dQ=0} \gtreqless 0 \quad \text{as} \quad [\partial c/\partial t - \theta c(P)/t] \gtreqless 0$$

for $q > 0$, $t > 0$.

Therefore, the value of t that minimizes the average cost of producing any quantity Q depends not only on the slope of c in t, but also on the returns to scale parameter θ. The interaction of these factors can be illustrated in an example.

Suppose there are only two inputs, capital K and labor H, where capital has a time invariant price p_k and labor's price varies directly with t, $p_H = tp_h$. We evaluate the bracketed term in (30) as

$$[\partial c/\partial t - \theta c(P)/t] = (\partial c/\partial p_k)(\partial p_k/\partial t) + (\partial c/\partial p_H)(\partial p_H/\partial t)$$
$$- \theta/t\ (p_k K + p_H H)$$

where K and H are the dual input demand functions from Lemma 2. Using Lemma 2 again and simplifying, we have

(31) $$[\partial c/\partial t - \theta c(P)/t] = p_h H - \theta[(p_k/t)K + p_h H]$$

since $\partial p_k/\partial t = 0$ and $p_H = tp_h$.

For $\theta = 1$, (31) is negative over all t and the firm choses $t = 1$. As θ increases (31) declines, implying that the firm is more likely to choose a small value of t when economies of scale are more pronounced. If we rewrite (31) for the more general case of $p_H = p_h(t)$, $p_h' > 0$, then

$$(32) \quad [\partial c/\partial t - \theta c(P)/t] = (\partial p_h/\partial t)H - \theta/t\ (p_K K + p_h H)$$

and we can derive the following condition

$$(33) \quad c(P)[S_H(\partial p_h/\partial t)/p_h - E_{Cq}/t] \gtreqless 0$$

where $S_H = p_h H/c(P)$ is labor's share of cost, and $\theta = E_{Cq}$ is the elasticity of cost with respect to q. Therefore the optimal t depends on the slope of labor's price in t as well as labor's share of cost and the scale properties of the rate function.

Equation (33) contains the results obtained by Betancourt and Clague in their propositions 2, 3, and 4, and by Winston for linearly homogeneous rate functions.[10] Yet (33) holds not only for homogenous rate functions of any degree as defined by (23), but also for homothetic rate functions. The reader may verify this result using the cost function dual to a homothetic production rate function,

$$C(q,P) = m(q)c(P) \text{ and } A(q,P) = q^{-1}\, t^{-1}\, m(q)c(P).$$

where m is a postive, increasing function. It then follows easily that

$$\left.\frac{dA}{dt}\right|_{dQ=0} = A[S_H(\partial p_h/\partial t)/p_h - E_{Cq}/t] \gtreqless 0$$

for the two input case.

VI. Conclusion

We have shown that dual cost and production functions exist for the flow-fund production technology derived from a process analysis of factory production. This duality includes the time utilization of the factors of production and asymmetric time patterns of payment of those factors. The usual results associated with the optimal utilization rate for linearly homogeneous rate functions can be extended to include homothetic rate functions.

The relationship between neoclassical production theory and the capital utilization literature is illuminated by the results of section V. If one ignores the time utilization problem, including the asymmetric variation in input prices with respect to time, and assumes a linearly homogeneous production function, then the flow-fund dual cost function collapses to the neoclassical cost function. Thus, the flow-fund production and cost relationships contain neoclassical production theory as a special case.

The flow-fund duality model encompasses the best attributes of neoclassical production theory, the mathematics of duality theory, and the time characteristics of capital utilization models. The neoclassical isoquant concepts remain useful and valid. The empirical vitality of duality theory is preserved.

The captial utilization problem is incorporated and extended to a wider range of technologies. In this way the unifying goal of this inquiry is achieved.

Furthermore, a general approach to production modeling is suggested. One begins with a process analysis of the system in question. This identifies the inputs and outputs concretely. Examination of the time pattern of input payments allows the construction of an appropriate cost function using duality theory. It is then a simple matter to derive the input demand equations for econometric estimation, given the appropriate data, or to estimate the cost function directly with the usual techiques.

This method also guarantees that the error for which Winston criticizes Shepherd, that "the representations of technology and input prices, and hence costs, that underlie and justify duality theory are either internally inconsistent or else applicable only to a firm that is, in very central ways, unlike any we know,"[11] is not repeated. One cannot properly apply the process analysis approach without unveiling the deficiency of neoclassical production duality analysis, "the failure to recognize that input flows to production differ in essential respects in their technological and ownership characteristics and that those differences are an integral part of the production process that must be captured either in its technological representation or in the representation of its prices and costs."[12] That "capture" has been effected here, in no small part, by following the instruction of Georgescu-Roegen:

"From all we know, only cost is a fact; the production functions are analytical fictions in a broader sense of the term than the formulas of the natural sciences. The latter are calculating devices, while the former are analytical similes which only help our Understanding to deal with a complex actuality pervaded by qualitative change. All the more necessary it is that these similes should be as faithful as Analysis can allow them to be."[13]

FOOTNOTES

1 Smith (1961) constructed a similar model based on the concept of "economic balance". Marsden, Pingry, and Whinston (1974) built a reaction model for chemical processes. Stewart (1980) and Cowing (1974) investigate steam electric power generation using an engineering approach.

2 This was pointed out by Georgescu-Roegen (1970), p. 1, some time ago.

3 In fact, J. M. Clark's (1923) work foreshadows much of this literature on capital utilization. This is one reason why Winston lauds the more intuitive and casual contributions to production theory under the blanket term neoclassical. Nevertheless, his condemnation of mathematical duality theory is misplaced. It is not the mathematics, but some well-worn interpretations of its applications that are at fault. Since these date from the marginalist period, we use the neoclassical label.

4 Some argue that the modern scientists' reliance on mathematics has blinded much research to the world of common sense. See, for instance, Barrett (1978) and Georgescu-Roegen (1971). A similar case can be made here: That economists in their zeal for developing a mathematical literature failed to incorporate some obvious characteristics of real production processes. Winston (1982), Chapter 6, attacks neoclassical duality models for this fault. At any rate, this tendency may have contributed to the belated attention given to questions of capital utilization and shiftwork.

[5] Georgescu-Roegen (1972), p. 284.

[6] Marsden, Pingry, and Whinston (1974), p. 136.

[7] It is a trivial exercize to show that

$$C(q,P) = \min_X \{P'X : f(X) \geq q,\ tq = 0\}$$

exists for f(X) satisfying Assumption 1 with input vector X and an appropriately modified price vector, $P = (p_1, \ldots p_m, tp_{m+1}, \ldots tp_n)'$, where all p_i and t are given.

[8] The Hessian matrix of the cost function with respect to the input prices evaluated at (q^*, P^*) is defined as

$$\nabla^2_{pp}C \equiv [\partial^2 C(q^*,P^*)/\partial P_i \partial P_j] \quad .$$

We know that

$$\partial X_i/\partial P_j \equiv [\partial X_i(q^*,P^*)/\partial P_j] = \nabla^2_{pp}C(q^*,P^*)$$

Where $\partial X_i/\partial P_j$ is the matrix of partial derivatives of the input demand functions with repsect to the input prices. Concavity of C in P, along with twice continuous differentiability of C with respect to P at (q^*, P^*), implies that the Hessian is a negative semidefinite matrix. Thus, we can find that

$$\partial X_i/\partial P_i \leq 0 \quad , \quad \text{for all } i,$$

or that the ith cost minimizing input demand function cannot slope upward with respect to its own price. Furthermore, twice continuous differentiability of C with respect to P at (q^*,P^*) implies that the hessian is a symmetric matrix, so the following symmetry conditions must hold

$$\partial X_i/\partial P_j = \partial X_j/\partial P_i, \text{ for all } i \text{ and } j.$$

[9] Diewert (1982), pp. 587-588.

[10] Betancourt and Clague (1981), pp. 18 and 26, and Winston (1982), p. 79.

[11] Winston (1982) p. 129.

[12] Winston (1982) pp. 131-132.

[13] Georgescu-Roegen (1972) p. 293.

REFERENCES

Barrett, William. The Illusion of Technique. Garden City, New York: Anchor/Doubleday, 1978.

Betancourt, Roger and Christopher Clague. Capital Utilization. Cambridge: Cambridge University Press, 1981.

Clark, J. M. The Economics of Overhead Costs. Chicago: University of Chicago Press, 1923.

Cowing, Thomas G. "Technical Change and Scale Economies in an Engineering Production Function: The Case of Steam Electric Power." Journal Of Industrial Economics, December, 1974.

Diewert, W. E. "Duality Approaches to Microeconomic Theory." Arrow, K. and N. Intrilligator, eds. Handbook of Mathematical Economics. Amsterdam: North-Holland Publishing Co., 1982.

Georgescu-Roegen, Nicholas. "The Economics of Production." Richard T. Ely Lecture, American Economic Review, May 1970, pp. 1-9.

__________. The Entropy Law and the Economic Process. Cambridge: Harvard University Press, 1971.

__________. "Process Analysis and the Neoclassical Theory of Production." American Journal of Agricultural Economics, May 1972, pp. 279-294.

Marris, Robin. The Economics of Capital Utilization. Cambridge: Cambridge University Press, 1964.

Marsden, James, David Pingry, and Andrew Whinston. "Engineering Foundations of Production Functions." Journal of Economic Theory, October 1974, pp. 124-139.

Smith, Vernon L. Investment and Production. Cambridge: Harvard University Press, 1961.

Stewart, John. "Plant Size, Plant Factor, and the Shape of the Average Cost Function in Electric Power Generation: A Non-Homogeneous Capital Approach." Bell Journal of Economics, Autumn 1979, pp. 549-565.

Winston, Gordon C. The Timing of Economic Activities. Cambridge: Cambridge University Press, 1982.