September 1982

FILE COPY
Do Not Remove From Files

A.E. Res. 82-28

THE TRANSLOG PRODUCTION FUNCTION: ITS PROPERTIES, ITS SEVERAL INTERPRETATIONS AND ESTIMATION PROBLEMS

by

Richard N. Boisvert

Department of Agricultural Economics
Cornell University Agricultural Experiment Station
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853

It is the policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

Table of Contents

The Translog Production Function: Its Properties, Its Several Interpretations and Estimation Problems

by

Richard N. Boisvert*

Although the theoretical importance of differing rates of substitution among productive inputs was recognized in the early 1930's, the first serious challenge to the simplifying assumptions embodied in the Cobb-Douglas and Leontief-type production functions came in the early 1960's. Until that time, much of the empirical work in production economics was at an aggregate level. Functions relating total production (or value added) in an economy to aggregate labor and capital inputs were offered in support of the marginal productivity theory of value: the data pointed to constant returns to scale and depicted a remarkable constancy of labor's share of output in the United States over time (Douglas, 1976).

Because there had been little empirical evidence to the contrary, economists, working on a variety of theoretical problems, were content with assuming either zero or unitary substitution elasticities among labor and capital. The simplistic assumptions were also made out of mathematical convenience or necessity, but as economists studied individual sectors of the economy, the seriousness of the limitations emerged. The starting point for the development of an alternative was "the empirical observation that the value added per unit of labor used within a given industry varies across countries with the wage rate" (Arrow, et al., 1961, p. 225). In addition to its obvious implications

*Richard N. Boisvert is an associate professor in the Department of Agricultural Economics, Cornell University, Ithaca, N.Y.

for factor shares and the functional distribution of income, the varying degrees of substitutability lead to reversals of factor intensities at different price ratios; the consequences for international trade are discussed by Minhas (1962).

Efforts to estimate the substitutability among productive inputs have intensified for two additional reasons. First, the desire to understand the processes of biased technical change and induced innovation in both developed and less-developed countries requires analytical models characterized by variable elasticities of substitution and capable of including many factors of production. Disaggregating labor and capital into subclasses can help refine policy implications of an analysis and is necessary when production processes are not separable between primary and intermediate factors (Binswanger, 1974).[1]

The second reason is the rapid change in relative prices between reproducible capital and labor and natural resource inputs since 1970. To develop policy measures for allocating natural resources one must understand the input intensities likely to arise from relative price

[1]The historical significance of the separability assumption is that it was used to justify estimating value added as a function of labor and capital alone. If the assumption were valid, no specification error would result from ignoring other inputs such as intermediate materials. Mathematically, separability deals with the appropriateness of being able to separate a function in many variables into subfunctions, each consisting of a smaller number of variables. Leontief (1947) demonstrated, for example, that a sufficient condition for two inputs, x_1 and x_2, to be functionally separable (weakly) from a third (x_3) is for the ratio of the marginal products of x_1 and x_2 to be independent of the level x_3. In this case, the production function for a product $Y = Y_1(x_1,x_2,x_3)$ can be written as $Y = Y_2[v_1(x_1,x_2)x_3] = Y_3[v, x_3]$, where $v = v_1(x_1, x_2)$.

changes, as well as the physical possibilities of substituting for nonrenewable or slowly regenerating natural resources whose supplies are extremely inelastic even at very high prices. This requires formal consideration of inputs other than aggregate labor and capital.

Several factors have contributed to difficulties encountered in efforts to estimate input substitutabilities. Perhaps the most important is the data limitation. Rarely does one find sufficient data on disaggregate input levels, prices, and output to specify appropriate analytical models either at the firm or industry levels. As the need for these kinds of data becomes more apparent, one can only hope that additional resources are devoted to collecting data in a more usable form.

A second factor has been the flexibility of the analytical models themselves. Although the CES production function, for example, accommodates elasticities of substitution different from zero or unity, they remain constant at all levels of input. The general applicability of the CES function has been restricted because of the nonlinear estimation problems and the necessity to choose among several alternative CES forms on the basis of functional separability (Uzawa, 1962).

These difficulties are in part responsible for the development of more flexible forms of production functions, the transcendental logarithmic (translog) production function (Christensen, Jorgenson and Lau, 1972) and generalized Leontief production function (Diewart, 1971). The "translog" form is the most widely used, perhaps because of its several possible interpretations and its mathematical similarity to the applications of Shephard's duality theory and translog cost functions.

Recent advances in econometric methods and the resolution of theoretical issues in measuring input substitution also help to explain its recent popularity.

Despite the frequent use of translog formulations, it is still quite difficult for graduate students and others to find a single document that describes in detail the properties and possible interpretations of the formulations. Even the earlier papers on the subject assume that the reader has a working knowledge of Taylor approximations in many variables, Shephard's Duality Theory, sufficient conditions for linear homogeneity of general functions, and the statistical relationships of mathematical models estimated in different algebraic forms.[2] Binswanger's (1975) discussion paper is the one exception in that he tries to demonstrate the simple mathematics of Shephard's Duality. His emphasis is on cost and profit functions and not on the translog production formulation.

Binswanger justifies the preoccupation with cost functions because of the ease with which it allows one to estimate Allen partial elasticities of substitution and their associated standard errors. However, there is a considerable loss of information in the sense that marginal productivities and the associated standard errors are difficult to derive. Also, the Allen partial elasticities of substitution are based on the assumptions of neoclassical markets. Although one may lose

---

[2]The most frequently cited reference on the properties of the translog production function is Christensen, Jorgenson, and Lau (1972), but the document is not readily accessible. Berndt and Christensen (1973) provide a useful summary of the function's properties but none of the derivations are provided.

some precision in interpreting statistical tests, the flexible production formulation may allow one to estimate other types of substitution elasticities that do not embody inherent behavioral assumptions.

The purpose of this paper is to derive in a systematic fashion the mathematical, economic and statistical relationships in the translog production formulations. The derivations are provided in detail, with the expectation that the paper can serve as a source document for graduate students or others trying to work with these models for the first time. At times, the algebra becomes tedious, but not unnecessarily so. A thorough knowledge of production economics is assumed.

The translog production function is discussed at the outset. A slight digression on the various definitions of the elasticity of substitution is needed to help distinguish between concepts which embody behavioral assumptions and those that do not.

The Translog Production Function

As with some other exponential functions, the translog production function is most often written in its logarithmic form, but for completeness, it is useful to write the function as

$$(1)\quad y = f(x_1,\ldots,x_n) = \alpha_0 \prod_{i=1}^{n} x_i^{\alpha_i} \prod_{i=1}^{n} x_i^{1/2[\sum_{j=1}^{n} \beta_{ij} \ln x_j]}$$

where

y = output;

α_0 = efficiency parameter;

x_j = input j; and

α_i and β_{ij} = unknown parameters.

Taking natural logarithms of both sides, one obtains the more familiar form[3]

$$(2)\quad \ln y = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j.$$

This algebraic formulation can be viewed in three ways: as an exact production function, as a second-order Taylor series approximation to a general, but unknown production function or as a second-order approximation to a CES production function. Each alternative

---

[3]Formulating the problem with the 1/2 in front of the quadratic expression is convenient in many of the derivations. However, some care is required when using parameter estimates obtained from statistical procedures directly in these derived expressions. The situation where this difference is important is highlighted throughout.

The equality of β_{ij} and β_{ji} for $i \neq j$ is assumed throughout to maintain consistency with Young's theorem of integrable functions (that the second cross partial derivative of the function with respect to i, then j, is equal to the second cross partial with respect to j, then i) (Berndt and Christensen, 1973).

These same authors mention briefly the possibility of including a technological index (A) into the translog function.

$$\ln y = \ln \alpha_0 + \alpha_A \ln A + \sum_{i=1}^{n} \alpha_i \ln x_i + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j$$

$$+ 1/2\, \beta_{AA} (\ln A)^2 + \sum_{i=1}^{n} \beta_{iA} \ln x_i \ln A.$$

Imposing Hicks-neutral technical change implies

$$\alpha_A = 1;\ \beta_{AA} = 0;\ \beta_{iA} = 0 \quad (\text{all } i).$$

By letting $\ln \alpha_0 = \ln \alpha_0' + \ln A$, the function assuming Hicks-neutral technical change is given by equation (2). These conditions are assumed throughout the theoretical sections for convenience of exposition.

interpretation may be more appropriate for some applications than others; therefore, each is worth discussing.[4]

Interpretation 1 - Exact Production Function

As an exact production function, equation (2) reduces to a Cobb-Douglas function in the case where all $\beta_{ij} = 0$. Thus, one immediate use offered by the translog production function is a straightforward test of the appropriateness of the maintained hypothesis embodied in the Cobb-Douglas function.[5]

More importantly, one must examine the function when at least one $\beta_{ij} \neq 0$, in which case it may or may not be well-behaved (e.g., output is monotonically increasing in all inputs and if the isoquants are convex).

To demonstrate these properties, it is convenient to begin with the production elasticities. From equation (2)

$$(3)\quad e_i = \frac{\partial \ln y}{\partial \ln x_i} = \alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j \qquad (i=1,\ldots,n).$$

The marginal products are

$$(4)\quad f_i = \frac{\partial y}{\partial x_i} = \left[\frac{\partial \ln y}{\partial \ln x_i}\right] [y/x_i] = [\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j][y/x_i].$$

---

[4]Having to treat three different interpretations separately is tedious. However, the literature to date appears to be somewhat imprecise in its discussions of the translog production function. One objective of the paper is to clarify the ambiguity.

[5]One test is the F-test described by Maddala (1977, p. 197) for testing linear restrictions in regression models. In this case, the restrictions are imposed by merely eliminating the quadratic terms. More is said about a general test of homogeneity below.

For finite levels of x_i, the marginal product of x_i can be positive for a range in values of x_j but can be negative if $\beta_{ij} > 0$ (all i, j) and $x_j \to 0$. Similarly, if there exists at least one $\beta_{ij} < 0$, $f_i < 0$ as $x_j \to \infty$. Thus, because monotonicity requires that for all i, $f_i > 0$, the translog function is not well-behaved globally.

The second direct and cross partial derivatives are obtained by applying the chain rule to equation (4). For all i and j,

$$(5)\quad f_{ii} = \frac{\partial y^2}{\partial x_i^2} = y[\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j] \, [\frac{-1}{x_i^2}]$$

$$+ \frac{1}{x_i}\left[y(\beta_{ii}/x_i) + \frac{y}{x_i}[\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j]^2 \right]$$

$$= \frac{y}{x_i^2}\left[-(\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j) + \beta_{ii} + [\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j]^2 \right]$$

$$= \frac{y}{x_i^2}\left[\beta_{ii} + (\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j - 1)\,(\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j) \right]$$

and

$$(6)\quad f_{ij} = \frac{\partial^2 y}{\partial x_i \partial x_j} = \frac{1}{x_i}\left[y(\beta_{ij})\,\frac{1}{x_j} + (\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j)(\alpha_j + \sum_{i=1}^{n} \beta_{ij} \ln x_i)\frac{y}{x_j} \right]$$

$$= \frac{y}{x_i x_j}\left[\beta_{ij} + (\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j)(\alpha_j + \sum_{i=1}^{n} \beta_{ij} \ln x_i) \right] .$$

The isoquants are strictly quasi-convex if the Bordered Hessian matrix

$$(7) \quad F = \begin{bmatrix} 0 & f_1 & f_2 & \cdots & f_n \\ f_1 & f_{11} & f_{12} & \cdots & f_{1n} \\ f_2 & f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_n & f_{n1} & \cdot & \cdots & f_{nn} \end{bmatrix}$$

is negative definite.[6] Because the values of the first and second partial derivatives vary with input levels, there is no guarantee that the isoquants are globally convex. However, in empirical research, "... there are regions in input space where these conditions are satisfied. If these conditions can be verified for each data point for any estimated translog function, the well-behaved region may be large enough to provide a good representation of the relevant production function" (Berndt and Christensen, 1973, p. 85). From equations (4) and (5), positive, but diminishing marginal productivity requires that $e_i > 0$ and $(e_i-1)e_i > \beta_{ii}$ if $B_{ii} < 0$. According to one definition (Ferguson, 1969), input i and input j are substitutes (complements) if $e_i e_j + \beta_{ij}$ is greater or less than zero (equation (6)).

Economists are also often interested in the rate at which output changes when all factors are changed by the same proportion. This output response is generally referred to as the economies of scale embodied in the production function. For a homogeneous production function, scale

---

[6]F is negative definite if the successive principle minors alternate in sign. Defining the k+1 principle minor by F_k, F is negative definite if $F_1 < 0$, $F_2 > 0$, $F_3 < 0$, ..., $(-1)^n F_n > 0$.

economies may be less than, equal to or greater than unity, but for a given function the "returns to scale" <u>are invariant</u> with respect to the initial input levels and are equal to the sum of the production elasticities. Frisch (1965) and Ferguson (1969, p. 81-83) establish that the "function coefficient" (the proportional change in output due to equal proportional changes in all inputs) is equal to the sum of the production elasticities for nonhomogeneous functions as well. The practical significance lies in the fact that for nonhomogeneous functions such as the translog function, the function coefficient is <u>not invariant</u> with initial input levels. From equation (3) the function coefficient (ε) is

$$(8)\quad \varepsilon = \sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \alpha_i + [\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_j].$$

Although there may be advantages in working with nonhomogeneous functions, one can derive sufficient conditions under which the translog function is homogeneous. In general, a function is homogeneous of degree h if

$$(9)\quad f(t\, x_1, \ldots, t\, x_n) = t^h(f(x_1, \ldots, x_n)).$$

If the function is written in logarithms, it is homogeneous of degree h if for any k

$$(10)\quad g(\ln x_1 + \ln k, \ldots, \ln x_n + \ln k) = h \ln k + g(\ln x_1, \ldots, \ln x_n).$$

Therefore, the translog function is homogeneous if one can find restrictions on α_i and β_{ij} <u>such that</u>

$$(11)\quad \ln y^+ = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i(\ln x_i + \ln k)$$

$$+ 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij}(\ln x_i + \ln k)(\ln x_j + \ln k)$$

$$= \ln y + h \ln k,$$

where ln y is defined as in equation (3). Expanding equation (11), one obtains

$$(12)\quad \ln y^{+} = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + \sum_{i=1}^{n} \alpha_i \ln k$$

$$+ 1/2 \, [\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j + \ln k \ln x_j + \ln k \ln x_i + (\ln k)^2]$$

$$= \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + 1/2 \, [\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j] \sum_{i=1}^{n} \alpha_i \ln k$$

$$+ 1/2[\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln k \ln x_j + \ln k \ln x_i + (\ln k)^2].$$

The desired expression in equation (11) obtains if one can restrict β_{ij} such that the last term in brackets in equation (12) vanishes. This occurs when

$$(13)\quad \sum_{i=1}^{n} \beta_{ij} = \sum_{j=1}^{n} \beta_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} = 0.$$ [7]

Thus, a sufficient condition for homogeneity is that the row and column sums of the coefficients on the quadratic expressions sum to zero. Furthermore, from equations (10), (12) and (13) it follows that the degree of homogeneity is given by

---

[7]Expanding this last term in brackets [ ] in (12) and rearranging, one has

$$(14)\quad \sum_{j=1}^{n} \ln x_j (\ln k \sum_{i=1}^{n} \beta_{ij}) + \sum_{i=1}^{n} \ln x_i (\ln k \sum_{j=1}^{n} \beta_{ij}) + [\ln k]^2 \, (\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij}).$$

Applying the results of (13) each term in ( ) goes to zero, the entire expression vanishes and (12) meets the conditions set out in (11).

$$(15) \quad \sum_{i=1}^{n} \alpha_i;$$

equation (2) is linear homogeneous if the sum of the α_i equals unity.

Another important result for the homogeneous case can be derived by applying the results of equation (12) to the function coefficient in equation (8). That is, if equation (13) holds, then the double sum in brackets in equation (8) vanishes and the function coefficient (or scale economies become independent of the input levels. This is similar to the results obtained in the Cobb-Douglas case. However, in the Cobb-Douglas case, the individual production elasticities are constant as well. For the homogeneous translog model, the individual production elasticities change as input levels change, but their sum remains constant.

Another practical implication of these results in equation (13) is that if one associates an error structure with equation (2) and estimates the parameters using ordinary least squares, it is possible to test the null hypothesis of homogeneity. Provided that the traditional assumptions about the general linear regression model can be made, the test of linear homogeneity in equation (2) involves the F-test for a general linear hypothesis. The derivation is cumbersome and is therefore relegated to Appendix A, where a special case of three inputs is considered.

The final characteristic of the production function that must be examined is the elasticity of substitution between inputs. There are several alternative definitions of the elasticity of substitution and in the case of the translog function, their derivations are complex computationally. A second interpretation of the function can facilitate this computational problem. Therefore, to avoid duplication, the discussion of substitution elasticities is deferred to the next section.

Interpretation 2 - Taylor Approximation

The second interpretation of the translog function is as a second order Taylor series approximation to an unspecified underlying production function. Allen (1937, pp. 456-58) demonstrates that if $h(q_1,q_2,\ldots,q_n)$ is a function in n variables and if $(r_1, r_2, \ldots, r_n)$ is a fixed point at which the n derivatives to the function exist, then

$$(16)\quad y = h(q_1 + r_1, q_2 + r_2, \ldots, q_n + r_n)$$

$$= h(r_1,\ldots,r_n) + \left[\sum_{i=1}^{n} q_i \frac{\partial h}{\partial q_i} \Big|_{r_1, \ldots, r_n} \right]$$

$$+ \left[\frac{1}{2!} \sum_{i=1}^{n} q_i^{\,2} \frac{\partial^2 h}{\partial q_i^{\,2}} \Big|_{r_i, \ldots, r_n} + \ldots + \frac{1}{n!} \sum_{i=1}^{n} \sum_{j=1}^{n} q_i q_j \frac{\partial^2 h}{\partial q_i \partial q_j} \Big|_{r_1, \ldots, r_n} \right]$$

$$+ \text{ higher order terms.}$$

These first three terms of the Taylor series expansion are exactly the translog function if one defines f as a logarithmic function[8]

$$\ln y^* = h(\ln x_1^*, \ldots, \ln x_n^*);$$

$$q_i = \ln x_i^*;\ r_i = \ln r_i;$$

$$\ln a_0 = h(\ln r_1, \ldots, \ln r_i);$$

$$a_i = \partial \ln y^*/\partial \ln x_i^*;$$

8Lower case letters rather than Greek symbols are used to indicate the parameters of this model in an effort to distinguish this formulation from equation (2).

$b_{ii} = \partial^2 \ln y^*/\partial(\ln x_i^*)^2$; and

$b_{ij} = \partial^2 \ln y^*/\partial \ln x_i^* \ln x_j^*$.

To apply this second-order approximation, one must select a specific point $(\ln r_1, \ldots, \ln r_n)$ around which the approximation is expanded. If one selects $r_1 = r_2 = \ldots r_n = 1$ so that $\ln r_1 = \ldots = \ln r_n = 0$, then equation (16) becomes

$$(17) \quad \ln y^* = \ln a_0 + \sum_{i=1}^{n} a_i \ln x_i^* + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \ln x_i^* \ln x_j^*$$

where

$\ln a_0 = f(0, \ldots, 0)$

$$a_i = \partial \ln y^*/\partial \ln x_i^* \Big|_{\ln x_i^* = 0, \text{ all } i}$$

$$b_{ij} = \partial \ln y^*/\partial \ln x_i^* \, \partial \ln x_j^* \Big|_{\ln x_i^*, \ln x_j^* = 0 \text{ for all } i \text{ and } j}.$$

Thus, the production elasticities and the logarithmic second derivatives for an unknown production function are approximated at the point $\ln x_i^* = 0$ for all i by the parameters of the function. These parameters in turn can be used to examine whether or not this unspecified production function is well-behaved at this particular point or to test the function for homogeneity, etc. The major drawback to interpreting the translog function as a second-order approximation is that the approximation to the true underlying function worsens as one moves away from the point around which the expansion is made.

There is, however, a practical reason for considering this interpretation in empirical analysis. <u>When estimated using conventional regression</u>

techniques, the important characteristics of the production function evaluated at the geometric means of y and x_i in equation (2) are identical to the ones in equation (17) because the Taylor expansion around zero is equivalent to scaling the data around the geometric mean. To see this let there be $t = 1, \ldots, T$ observations on the variables y_t and x_{it} so that equation (2) can be written as in Appendix A.

$$(18)\quad \ln y_t = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_{it} + \sum_{i=1}^{n} \gamma_{ii}(\ln x_{it})^2 + \sum_{j>i}^{n}\sum^{n} \beta_{ij} \ln x_{it} \ln x_{jt} + u_t .$$

where $\gamma_{ii} = 1/2\ \beta_{ii}$ and u_t is an error term. To establish the equivalence in the characteristics of the functions, one must verify that such a scaling procedure implies that $E \ln x_i^* = 0$ and establish a relationship between the estimated parameters.[9]

The first task is to define $\ln x_{it}^*$

$$\ln x_i^* = \ln \left[x_{it} / (\prod_{t=1}^{T} x_{it})^{1/T} \right] = \left[\ln x_{it} - 1/T \sum_{t=1}^{T} \ln x_{it} \right] \quad ;$$

[9]As mentioned in footnote 1, the estimating form of the equation internalizes the 1/2 in front of the quadratic expression in equation (2) and does not distinguish between $\ln x_{it} \ln x_{jt}$ and $\ln x_{jt} \ln x_{it}$. The estimating form of (17) is

$$\ln y^* = \ln a_0 + \sum_{i=1}^{n} a_i \ln x_{it}^* + \sum_{i=1}^{n} g_{ii}(\ln x_i^*)^2 + \sum_{i<j}^{n}\sum^{n} b_{ij} \ln x_{it}^* \ln x_{jt}^* + \varepsilon_t$$

where

$g_{ii} = 1/2\ b_{ii}$ and ε_t is an error term.

and derive its expected value

$$E \ln x_i^* = E\left[\ln x_{it} - 1/T \sum_{t=1}^{T} \ln x_{it}\right]$$

$$= \frac{1}{T} \sum_{t=1}^{T} \left[\ln x_{it} - \frac{1}{T} \sum_{t=1}^{T} \ln x_{it}\right]$$

$$= 1/T \sum_{t=1}^{T} \ln x_{it} - \frac{1}{T^2}\, T \sum_{t=1}^{T} \ln x_{it} = 0.$$

The second part is more difficult to establish, but algebraically, the procedure is to expand the estimating form of equation (17) (c.f. footnote 9). For convenience, let

$$\tilde{\ln}\, x_i = \frac{1}{T} \sum_{t=1}^{T} \ln x_{it};\ \tilde{\ln}\, y = \frac{1}{T} \sum_{t=1}^{T} \ln y_t\ .$$

Scaling the equation from footnote 9

$$(19)\quad \ln y_t - \tilde{\ln}\, y = \ln a_0 + \left[\sum_{i=1}^{n} a_i(\ln x_{it} - \tilde{\ln}\, x_i)\right] + \{\sum_{i=1}^{n} g_{ii}(\ln x_{it} - \tilde{\ln}\, x_i)^2$$

$$+ \sum_{i<j}^{n} \sum^{n} b_{ij}(\ln x_{it} - \tilde{\ln}\, x_i)(\ln x_{jt} - \tilde{\ln}\, x_j)\} + e_t.$$

Expanding the term in square brackets, one has

$$(20)\quad \sum_{i=1}^{n} a_i(\ln x_{it} - \tilde{\ln}\, x_i) = \sum_{i=1}^{n} a_i \ln x_{it} - \sum_{i=1}^{n} a_i\ \tilde{\ln}\, x_i;$$

and expanding the term in { }, one has

$$(21)\quad \sum_{i=1}^{n} g_{ii}(\ln x_{it}-\tilde{\ln}\, x_i)^2 + \sum_{i<j}^{n}\sum^{n} b_{ij}(\ln x_{it}-\tilde{\ln}\, x_i)(\ln x_{jt}-\tilde{\ln}\, x_j)$$

$$= \sum_{i=1}^{n} g_{ii}(\ln x_{it})^2 - 2\sum_{i=1}^{n} g_{ii}\ln x_{it}\ \tilde{\ln}\, x_i + \sum_{i=1}^{n} g_{ii}(\tilde{\ln}\, x_i)^2$$

$$+ \sum_{i<j}^{n}\sum^{n} b_{ij}\ln x_{it}\ln x_{jt} - \sum_{i<j}^{n}\sum^{n} b_{ij}\ln x_{it}(\tilde{\ln}\, x_j)$$

$$- \sum_{i<j}^{n}\sum^{n} b_{ij}(\tilde{\ln}\, x_i)(\ln x_{jt}) + \sum_{i<j}^{n}\sum^{n} b_{ij}(\tilde{\ln}\, x_{it})(\tilde{\ln}\, x_{jt}).$$

Substituting back into (19) and collecting some common terms one has

$$(22)\quad \ln y_t = \ln a_0 + \tilde{\ln}\, y - \sum_{i=1}^{n} a_i\ \tilde{\ln}\, x_i + \sum_{i<j}^{n}\sum^{n} b_{ij}(\tilde{\ln}\, x_i)(\tilde{\ln}\, x_j)$$

$$+ \sum_{i=1}^{n} g_{ii}(\tilde{\ln}\, x_i)^2 + \sum_{i=1}^{n}(a_i - \sum_{j\neq i}^{n} b_{ij}\ \tilde{\ln}\, x_j - 2g_{ii}\ \tilde{\ln}\, x_i)\ln x_{it}$$

$$+ \sum_{i=1}^{n} g_{ii}(\ln x_{it})^2 + \sum_{i<j}^{n}\sum^{n} b_{ij}(\ln x_{it})(\ln x_{jt}) + e_t.$$

Thus, the two models in equations (18) and (22) are equivalent if we let

$$(23)\quad \ln \alpha_0 = \ln a_0 + \tilde{\ln}\, y - \sum_{i=1}^{n} a_i\ \ln x_i + \sum_{i<j}^{n}\sum^{n} b_{ij}(\tilde{\ln}\, x_i)(\tilde{\ln}\, x_j)$$

$$+ \sum_{i=1}^{n} g_{ii}(\tilde{\ln}\, x_{it})^2;\ g_{ii} = \gamma_{ii};\ b_{ij} = \beta_{ij};\ \text{and}$$

$$\alpha_i = (a_i - \sum_{j\neq i}^{n} b_{ij}\ \tilde{\ln}\, x_j - 2g_{ii}\ \tilde{\ln}\, x_i).$$

Because the coefficients on the squared and cross-product terms are identical,

$$(24)\quad \alpha_i = a_i - \sum_{j\neq i}^{n} \beta_{ij}\ \tilde{\ln}\, x_j + 2\gamma_{ii}\ \tilde{\ln}\, x_i,\ \text{and}$$

$$(25)\quad a_i = \alpha_i + \sum_{j\neq i}^{n} \beta_{ij} \tilde{\ln} x_j + 2\gamma_{ii} \tilde{\ln} x_i .$$

Recalling the relationship between γ_{ii} and β_{ii} and g_{ii} and b_{ii} in the different forms of the models, this demonstrates that the production elasticities estimated directly from the scaled model (equation 17) are identical to the production elasticities of the unscaled model (equation 2) evaluated at the geometric mean.

In summary, it has been shown that there is a direct correspondence between these two interpretations of the translog production function at the point around which the second model is expanded. Through a similar procedure, one can establish that the same relationship holds for other values of x and y as well. That is, if one were to expand equation (17) around any other interesting point (say the arithmetic mean of x_i), the relationships in equations (24) and (25) would hold by making the appropriate substitution for $\tilde{\ln} x_j$. The practical implications of this correspondence is that one can derive the first and second order partial derivatives and other production relationships for a particular point on the unscaled function by estimating the scaled model only. This simplifies the calculation of marginal products and second direct and cross partial derivatives a great deal.[10] More importantly, the correspondence

---

[10]That is, if the data are scaled at the geometric mean (both x and y) a_i, b_{ij} and b_{ii} are interpreted directly as the first and second order logarithmic derivatives. Calculating f_i and f_{ij} and f_{ii}, from expressions equivalent to equations (4), (5) and (6) becomes easy because $\ln x_i = 0$ as a result of the scaling. However, y and x in these equations are at the geometric means of the variables in unscaled units.

simplifies the calculations of various measures of the elasticity of substitution between inputs.[11]

Elasticities of Factor Substitution

The discussion of factor substitution possibilities and their measurement dates to the 1930's with the work of Robinson (1933) and Hicks (1970). Initial discussions were limited to the case of two inputs and the elasticity of substitution was defined as: the proportional change in the input ratio due to a proportional change in the marginal technical rate of substitution, output held constant.

Allen (1938) is credited with the development of a measure of a partial elasticity of substitution between any two inputs in an n-factor production system. It is related to the demand for factors under the assumption of competitive markets and profit maximizing behavior. The Allen partial is defined as:

> the effect on the quantity demanded of one factor of a change in the price of another factor, where the partial derivative is taken holding output and other factor prices constant (Sato and Koizumi, 1973, p. 47).

For two factors, i and j, the Allen partial for equation (2) is given by

$$(26) \quad \sigma_{ij} = \frac{\sum_{j=1}^{n} x_j f_j}{x_i x_j} \bullet \frac{F_{ij}}{F}$$

---

[11]The translog function can also be interpreted as a Taylor Series Approximation to a CES production function. Kmenta (1967) and Griliches and Ringstad (1971) have used a two input version, but the parameters of a more general form in many variables are underidentified. Thus, the derivation is relegated to Appendix C and is primarily of academic interest.

where F is the determinant of the bordered Hessian matrix in (7) and F_{ij} is the cofactor of f_{ij} and F. While Ferguson (1969) and others have demonstrated that the Allen partials for the C-D function are equal to unity for all i and j, for the translog function (2) σ_{ij} can be positive, negative or zero. The only restrictions on the values are that

$$(27) \quad \sum_{j=1}^{n} k_j \; \sigma_{ij} = 0; \quad \text{where } k_j = \frac{x_j f_j}{\sum_{j=1}^{n} x_j f_j} \; ;$$

for linear homogeneous functions. For functions that are not linear homogeneous, Ferguson (1969, pp. 180-85) derives the relationship between σ_{ij} and input demands under cost minimization assumptions.

The original concept of the elasticity between factors in the two-factor case has also been extended to the n-factor case. It is called the direct elasticity of factor substitution (DES) and is defined as the

> ratio between a percentage change in the factor proportion and a percentage change in the marginal rate of substitution given all other factors [and output] (Sato and Koizumi, p. 54).

In mathematical terms, the direct elasticity of substitution is given by

$$(28) \quad e_{ij} = \frac{d(x_j/x_i)}{x_j/x_i} \div \frac{d(f_i/f_j)}{f_i/f_j} \quad .$$

Stated in this way, it is the generalization of the two-factor elasticity of substitution discussed by Ferguson (1969, p. 91). Exactly why this concept of the direct elasticity of substitution has only recently been applied in the n-factor case is unclear. One rationale is that it still makes it difficult to classify inputs. This, however, seems to be a minor problem. The Allen partial, one alternative which is so dependent on the assumption of linear homogeneity and economic rationality narrowly defined, has its own problems. One might argue that what is really important is the relative curvature of the isoquants as measured by the direct elasticity of substitution from which response to changing prices can always be obtained anyway. (There are other elasticity concepts as well and Sato and Koizumi (1973) do as good a job as anyone in establishing the formal relationships among them.)

The computation of the DES for the translog function is complicated algebraically. Therefore, the derivation is relegated to Appendix C. It is also demonstrated in this appendix that computations are simplified in the case of interpreting the translog function as a second-order approximation to an unknown function (e.g., the data are scaled around the geometric mean).

Input Separability

Separability of inputs is an issue that is often discussed in aggregate production analysis. Discussions of separability relate to the internal structure of functions and whether a function of many arguments can be separated into subfunctions. There are essentially two reasons for this interest. If functions are separable in some groups of inputs, production decisions and relative factor intensities can be optimized within each subset, and then optimal factor intensities can be obtained by holding fixed the within-subset intensities and optimizing the between-subset intensities (Berndt and Christensen, 1973). If a production function is separable in several groups of inputs, then the inputs within a subset can be aggregated into a single composite input.[12]

---

[12]Leontief (1947) and Solow (1954-55) were among the first economists to discuss this issue.

The practical implications of the interest in separability is to enable one to study complex production relations in a piecemeal fashion. This often simplifies statistical analysis and is often necessary because of the lack of data.

To begin the discussion of separability, consider (as Berndt and Christensen do), a twice differentiable, strictly concave homothetic production function with strictly positive marginal products.[13]

(29) $Y = f(x) = f(x_1, x_2, \ldots, x_n)$.

The set of inputs $N = [1, \ldots, n]$ is partitioned into r mutually exclusive and exhaustive subsets $[N_1, \ldots, N_r]$, a partition denoted R.

The production function f(x) is said to be weakly separable with respect to the partition R if the marginal rate of substitution between two inputs x_i and x_j from any subset $N_s (s=1, \ldots, r)$ is independent of the quantities outside N_s. That is,

(30) $\frac{\partial}{\partial x_k} \left(\frac{f_i}{f_j}\right) = 0$, for all i, j εN_s and k $\not\varepsilon N_s$.[14]

---

[13]A function f(X) is homothetic if it can be written as h(g(X)) where h is monotonic and g is homogeneous of degree 1.

[14]Strong separability is when the marginal rate of substitution between input i εN_s and input j εN_t is independent of inputs outside N_s and N_t. That is,

$$\frac{\partial}{\partial x_k} \left(\frac{f_i}{f_j}\right) = 0, \text{ for all } i \;\varepsilon\; N_s,\; j \;\varepsilon\; N_t,\; k \not\varepsilon\; N_s \cup N_t.$$

This condition implies weak separability but the reverse implication applies only in the case of two subsets.

Performing the differentiation implied in equation (30) gives an alternative separability condition:

(31) $f_j f_{ik} - f_i f_{jk} = 0.$

The major result from functional separability is given by the following theorem:

> Weak separability with respect to the partition R is necessary and sufficient for f(X) to be of the form $f(X^1, X^2, \ldots, X^r)$ where X^s is a function of the elements of N_s only.

This theorem was proven in a less general form by Leontief, and as Solow points out, X^s is a consistent aggregate index of inputs in N_s. That is, for production purposes, weak separability implies that any pattern of inputs in N_s is equivalent as long as they yield the same index value of X^s (e.g. X^s is the output of a sub-production function in the inputs in N_s). Thus, it follows that a consistent set of aggregate inputs exists if and only if the inputs are weakly separable from the others not in N_s.[15]

Berndt and Christensen (1973) demonstrate formally that weak separability implies equality of the Allen partial elasticity of substitution,

(32) $\sigma_{ik} = \sigma_{jk}$ $(i, j \in N_s, k \notin N_s)$;

but the best intuitive explanation of separability is given by Humphrey and Moroney (1975). Suppose that the usage of x_i and x_j is held constant

---

[15]Strong separability with respect to the partition R is necessary and sufficient for F(X) to be of the form $f(X^1 + X^2 + \ldots + X^r)$, where X^s is a function only of the inputs in N_s. This and the theorem in the text are established by Goldman and Uzawa, 1964.

and use of x_k increases. If this renders x_i and x_j more effective at the margin and their individual effectiveness is changed by exactly the same amount, then x_i and x_j are functionally separate from x_k. Marginal products are shifted in the same proportion, observationally the same as a Hick's neutral technical change or a change in efficiency. Thus, if x_k is an explicit third input, it is reasonable that if the marginal products for x_i and x_j are shifted vertically in the same proportion, x_k bears the same equally close substitution or complementarity relationship to both input i and j (equation (32) holds).

An Example of an Aggregate Index

Perhaps the best way to understand the implications of separability is to examine its implications for a separable function such as the Cobb-Douglas production function of the form

(33) $Y = L^{\alpha} C_1{}^{\beta} C_2{}^{\gamma};\ \alpha + \beta + \gamma = 1;\ \alpha;\ \beta,\ \gamma \geq 0.$

The marginal rate of substitution between the two types of capital, C_1 and C_2 is given by

(34) $$\frac{f_{C_1}}{f_{C_2}} = \frac{\beta\ Y/C_1}{\gamma\ Y/C_2} = \frac{\beta C_2}{\gamma C_1}\ ;\text{ and}$$

it is independent of the level of the other input.

Thus, according to the separability theorem, one should be able to find a consistent aggregate for capital and write

(35) $Y = f(L, K)$

where

(36) $K = F(C_1, C_2)$.

This is true if we let

$$(37)\quad K = C_1^{\frac{\beta}{\beta+\gamma}} C_2^{\frac{\gamma}{\beta+\gamma}}$$

and

$$(38)\quad Y = L^{\alpha} K^{\beta+\gamma}.$$

By substituting (37) into (38) we obtain the original equation (33). A logical extension of this argument provides the rationale for treating value added as a function of capital and labor only and ignoring other inputs as has been done many times. If labor and capital are separable from other inputs, then value added (the combined output of labor and capital) can be written as a function of labor and capital alone. (This illustration is due to Solow, 1954-55.)

Conditions for Separability in the Translog Production Function

The situation for the Cobb-Douglas function is in contrast to the situation in the translog case. Let us examine the translog function.

$$(39)\quad \ln y = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j .$$

If we consider the case where any inputs i and j are functionally separable from a third k, then we require

$$(40)\quad \frac{\partial y}{\partial x_i} \frac{\partial^2 y}{\partial x_j \partial x_k} - \frac{\partial y}{\partial x_j} \frac{\partial^2 y}{\partial x_i \partial x_k} = 0 .$$

Evaluating this expression by substituting from equations (4) and (6)

$$(41)\quad \frac{y}{x_i} [\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j] [\frac{y}{x_j x_k} \{\beta_{kj} + (\alpha_k + \sum_{j=1}^{n} \beta_{kj} \ln x_j)(\alpha_j + \sum_{i=1}^{n} \beta_{ij} \ln x_i)\}]$$

$$- \frac{y}{x_j} [\alpha_j + \sum_{i=1}^{n} \beta_{ij} \ln x_i] [\frac{y}{x_i x_k} \{\beta_{ik} + (\alpha_i + \sum_{j=1}^{n} \beta_{ij} \ln x_j)(\alpha_k + \sum_{j=1}^{n} \beta_{kj} \ln x_j)\}]$$

$$= 0.$$

This expression (41) simplifies to (recalling that e_i is defined as $\partial \ln y/\partial \ln x_i$):

(42) $\frac{y^2}{x_i x_j x_k} e_i[\beta_{kj} + e_k(e_j)] - e_j[\beta_{ik} + e_i\ e_k] = 0.$

Because $y^2/x_i x_j x_k > 0$, we can divide through by it and obtain

(43) $e_i\ \beta_{kj} + e_i\ e_k\ e_j - e_j\ \beta_{ik} - e_i\ e_j\ e_k = 0$

or the conditions of functional separability hold, if and only if

(44) $e_i\ \beta_{kj} - e_j\ \beta_{ik} = 0.$

For a well-behaved production function, we require that $e_i > 0$. Therefore, if the function is separable and if $\beta_{kj} = 0$, then $\beta_{ik} = 0$. If, however, $\beta_{jk} \neq 0$ and $\beta_{ik} \neq 0$, then we can expand (43) and find that i and j are separable from k if and only if

(45) $\beta_{kj}(\alpha_i + \sum_{m=1}^{n} \beta_{im} \ln x_m) - \beta_{ik}(\alpha_j + \sum_{m=1}^{n} \beta_{jm} \ln x_m) = 0$

or (because $\beta_{kj} = \beta_{jk}$)

(46) $\alpha_i\ \beta_{jk} - \alpha_j\ \beta_{ik} + \sum_{i=1}^{m} (\beta_{im}\ \beta_{jk} - \beta_{ik}\ \beta_{jm}) \ln x_m = 0.$

Thus, the necessary and sufficient conditions for global separability (independence of the x's) we require

(47) $\alpha_i\ \beta_{jk} - \alpha_j\ \beta_{ik} = 0;$

and

(48) $\beta_{im}\ \beta_{jk} - \beta_{ik}\ \beta_{jm} = 0 \quad (m=1,\ldots,n).$

If β_{jk} and β_{jm} are not equal to zero, we have

(49) $\frac{\alpha_i}{\alpha_j} = \frac{\beta_{ik}}{\beta_{kj}} = \frac{\beta_{im}}{\beta_{jm}} \ (m=1,\ldots,n).$

To the author's knowledge, there has been several attempts to test for separability in the translog production function, including Berndt and

Christensen, 1973, and Humphrey and Moroney, 1975. However, Berndt and Christensen are the only ones who attempt to then aggregate groups of separable inputs. Other analysts have tested for input separability when using a cost function to estimate elasticities of input substitution (e.g. Berndt and Wood, 1975). In cases where tests of separability of inputs were rejected, the authors have usually gone on to examine the substitution between labor and capital and other inputs such as natural resources, energy and intermediate inputs.

Practical Problems of Estimating Translog Production Function

To delineate the statistical tests in Appendix A, it was assumed that the translog production function was estimated as a single equation in the form:

$$(50)\quad \ln y_t = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_{it} + \sum_{i=1}^{n} \gamma_{ii}(\ln x_{it})^2 + \sum_{i<j}\sum \beta_{ij} \ln x_{it} \ln x_{jt} + u_t$$

where

$t = 1,\ldots,T$, the number of observations;

u_t = an error term;

$\gamma_{ii} = 1/2\ \beta_{ii}$ from equation (2).

Two serious problems in estimating this function by single equation methods are readily apparent. First, as the number of factors of production is increased, the number of parameters to be estimated increases rapidly. Because the additional terms are squares and cross products of the variables, multicollinearity is a difficult problem. Second, data may be limiting.

One potential solution to multicollinearity is to remove selectively those squared or cross product terms whose t-ratios are below a certain critical value. This strategy could ultimately destroy the flexibility in the relationships among inputs and after all, this flexibility is a major reason for considering the function in the first place. Shih, Hushak and Rask (1977) utilize this strategy quite effectively. Vinod (1972), on the other hand, proposes a functional form that differs from equation (2) only in the fact that all squared terms are eliminated. This preserves much of the flexibility of the function and undoubtedly does mitigate the multicollinearity problem to some extent. However, there seems to be little economic rationale for eliminating the squared terms a priori.

There are two alternative procedures for estimating the model if one is willing to assume linear homogeneity and profit maximizing behavior. The first is provided in Appendix A, whereby single equation methods are used to estimate a model in which the sufficient homogeneity conditions are imposed by transforming the original variables.

The second involves the assumption of profit maximization. To keep the algebra manageable, consider three inputs, the minimum number for which the problem is interesting. Let the production function be

$$(51) \quad \ln y = \alpha_0 + \sum_{i=1}^{3} \alpha_i \ln x_i + 1/2 \sum_{i=1}^{3} \sum_{j=1}^{3} \beta_{ij} \ln x_i \ln x_j .$$

Assume that the entrepreneur is a price taker in both factor and product markets and attempts to maximize profits (p = price of output; and r_i = price of input i)

$$(52) \quad \max \Pi = py - \sum_{i=1}^{3} r_i x_i .$$

Using the general expressions above for marginal products in a translog function, the first order conditions for profit maximization can be written as

$$(53)\quad \frac{\partial y}{\partial x_i} = \frac{\partial \ln y}{\partial \ln x_i} \bullet \frac{y}{x_i} = \left[\alpha_i + \sum_{j=1}^{3} \beta_{ij} \ln x_j\right] y/x_i = \frac{r_i}{p} \quad (i=1,2,3);$$

or

$$(54)\quad e_i = \frac{r_i x_i}{py} = \alpha_i + \sum_{j=1}^{3} \beta_{ij} \ln x_j \quad (i=1,2,3)$$

e_i = value of the ith factor relative to the value of output and in this case, also the production elasticity.

Furthermore, if there is long run competitive equilibrium and/or production is subject to constant returns to scale,

$$(55)\quad \sum_{i=1}^{3} e_i = 1.$$

If one can justify writing equation (54) in stochastic form (attributing the error to "mistakes" in trying to satisfy the first order conditions or adjustment lags), one can write[16]

$$(56a)\quad e_1 = \alpha_1 + \beta_{11} \ln x_1 + \beta_{12} \ln x_2 + \beta_{13} \ln x_3 + U_1$$

$$(56b)\quad e_2 = \alpha_2 + \beta_{21} \ln x_1 + \beta_{22} \ln x_2 + \beta_{23} \ln x_3 + U_2$$

$$(56c)\quad e_3 = \alpha_3 + \beta_{31} \ln x_1 + \beta_{23} \ln x_2 + \beta_{33} \ln x_3 + U_3.$$

If one assumes that the function is linear homogeneous and the symmetry restrictions are applied, the sum of the e_i add to unity and only two of the three equations are linearly independent. Thus, the parameter

[16]Now that the production function is not Cobb-Douglas, the identification problem with the "mongrel" combination of first order conditions is not a problem (Zellner, Kmenta and Dreze, 1967). From this one standpoint, the parameters are identifiable.

estimates of any two of the equations exactly identify the parameters of the production function.

From (56a) and (56b)

(57) $\alpha_3 = 1 - \alpha_2 - \alpha_1 \Rightarrow \hat{\alpha}_3 = 1 - \hat{\alpha}_2 - \hat{\alpha}_1$

Similarly, because of the symmetry restrictions

(58) $\hat{\beta}_{ij} = \hat{\beta}_{ji}$; $(i \neq j)$; $i,j = 1,2,3$

and because of the restriction on the row and column sums of the β matrix,

(59) $\hat{\beta}_{ij} = -(\hat{\beta}_{ii} + \hat{\beta}_{ik})$; for i,j and $k = 1,2,3$.

Substituting (58) into (56) we have

(60a) $e_1 = \alpha_1 + \beta_{11} \ln x_1 + \beta_{12} \ln x_2 + \beta_{13} \ln x_3 + U_1$

(60b) $e_2 = \alpha_2 + \beta_{12} \ln x_1 + \beta_{22} \ln x_2 + \beta_{23} \ln x_3 + U_2$

(60c) $e_3 = \alpha_3 + \beta_{13} \ln x_1 + \beta_{23} \ln x_2 + \beta_{33} \ln x_3 + U_3$

We know from (59) that

(61) $\beta_{13} = -\beta_{11} - \beta_{12}$; $\beta_{23} = -\beta_{12} - \beta_{22}$.

Therefore, we have by substituting into (60a) and (60b)

(61a) $e_1 = \dfrac{r_1 x_1}{py} = \alpha_1 + \beta_{11} \ln x_1 + \beta_{12} \ln x_2 - (\beta_{11} + \beta_{12}) \ln x_3 + U_1$

(61b) $e_2 = \dfrac{r_1 x_2}{py} = \alpha_2 + \beta_{12} \ln x_1 + \beta_{22} \ln x_2 - (\beta_{12} + \beta_{22}) \ln x_3 + U_2$.

These are the estimating equations; β_{33} and α_3 are calculated <u>ex post</u> by substituting the parameter estimates into equations (57) and (59).

This, however, does not mean that the estimation is easy. The first problem is that both price and quantity information are needed for each of the inputs. Second, as Humphrey and Moroney (1975, pp. 65-66) suggest, one would expect random deviations from profit maximization to affect all markets. One would hypothesize that U_1 and U_2 are correlated, implying the need for some sort of two-stage estimation. However, the estimates obtained by applying ordinary least squares or Zellner two-stage least squares depend on which two equations from (60a)-(60c) are chosen. Maximum-likelihood estimates would be independent of which equations were selected. Both Kmenta and Gilbert (1968), through Monte Carlo experiments, and Rubel, through formal proof, have demonstrated that iterative Zellner (IZ) estimates and maximum-likelihood methods are computationally equivalent. Accordingly, one must estimate parameters by applying the IZ method. Assuming that the elements of the regressor matrix are predetermined variables, application of IZ produces consistent and asymptotically efficient estimates of the parameters. Because all of the remaining parameters are linear combinations of these estimates, they also have these desirable asymptotic properties.

Some Concluding Comments

The purpose of this report is to derive the mathematical properties of the translog production function, discuss its several interpretations and describe the various approaches to estimating the function statistically. Emphasis is placed on three separate interpretations: a) as an exact production function; b) as an approximation to the CES function; and c) as a second order approximation to an unknown function. Although the report contains no empirical application, the presentation has several important implications for empirical analysis.

Perhaps the most important conclusion is that one encounters numerous computational difficulties and additional data requirements in order to take advantage of the function's flexibility. Therefore, before choosing to utilize the translog function, one must have a compelling *a priori* reason for believing that such flexibility is necessary to represent the production technology accurately. Even then, use of the function may not be justifiable if estimates of elasticities of substitution are less important to the analysis than are estimates of marginal products, scale elasticities or input demand relations. This is particularly true given that the Cobb-Douglas (C-D) function is a special case of the translog function. Historically, the Cobb-Douglas function has been used extensively in the literature to study both micro and more aggregate production problems. Since the performance of the C-D function has been extremely good from a statistical point of view, one might expect little to be gained from a more complex structure unless the motivation underlying the research is a test of the maintained hypothesis embodied in the C-D function.

This is not to argue that one should never employ this type of flexible production function in empirical analysis. However, in addition to the more demanding data requirements, there are several reasons why flexible production functions such as the translog function should be applied with extreme caution. The first reason has to do with the function's several possible interpretations. As pointed out above, the translog function can be viewed as an approximation to a CES production function and, except in the case of two inputs, the parameters of the underlying CES function are overidentified. Therefore, in this author's opinion, this particular interpretation of the translog function is of little more than academic interest.

The translog's interpretation as an exact production is perhaps the most appealing, but it is troublesome as well. The problems encountered in obtaining reliable estimates of the parameters of the function are difficult, if not impossible to resolve. Equally as important is the fact that the function is not well-behaved over the entire range of possible input levels. Although Berndt and Christensen argue that an estimated translog function may contain a large enough well-behaved region to provide a good representation of the production surface, it is unclear to this author how one can determine the size of this region. An obvious first step would be to check conditions at each data point, but it is doubtful that this is sufficient to guarantee that the function is well-behaved throughout the extremes of the data.

By far the most widely used interpretation of the translog productin function is as a second-order approximation to the true but unknown production system. In these cases, the performance of the models from a statistical point of view seemed much better than when treating it as an exact production function (e.g., Wyzan, 1981; Shih, et al., 1977; Dunne, 1981; Dunne and Boisvert, 1982). In particular, the t-ratios on many of the coefficients improved tremendously because the data were scaled and the coefficients on the log-linear terms carried an interpretation of production elasticities (see equation (25)) at the geometric means of the data. In cases where the translog model represented a marginal improvement over an acceptable C-D specification, it is hardly surprising that the production elasticity estimates or the level of confidence one has in them, would differ substantially at the point of geometric means.

The improved results obtained by scaling the data could mean that one has merely approximated a true underlying C-D function or that the translog model is a satisfactory approximation to the function at that point. In the latter case, one must be concerned about the range in the input levels over which the approximation remains satisfactory. However, even when the quadratic approximation to the function is adequate, the results, such as marginal products, production elasticities, scale economies etc. derived from it are only first-order approximations. Thus, as Theil (1980) suggests, the statistical analysis of the translog production function is a classic example of the tradeoff between the quality of approximation achieved by the specification and the statistical quality of the estimates of the parameters of the specification. "The approximation is usually satisfactory when the independent variables vary very little, but precise parameter estimates require adequate variation in these variables" (p. 151).

In conclusion, it may seem strange to end such a detailed treatment of the translog production function on what appears to be an extremely pessimistic note. In reading the last several paragraphs, one could obtain the distinct impression that the translog production system should never be used. However, these remarks are not for this purpose at all. Rather they are to encourage production economists to understand all flexible functional forms before using them in empirical analysis and ask the tough questions before accepting results derived from them. The literature abounds with applications of both the translog production and cost systems in which the theoretical and statistical problems raised above were not addressed adequately. There is also other literature which suggests that because the translog cost and productin systems are not self-dual, then at best one of the estimates of common production parameters that can be derived from both using the same data is wrong (Burgess, 1975). What is needed is a more exhaustive investigation into the conditions under which all flexibile production functions and or cost function (through application of duality) provide acceptable estimates of underlying production and input demand systems.

References

Allen, R. G. D. Mathematical Analysis for Economists. New York: St. Martin's Press, 1938.

Arrow, K. J., H. B. Chenery, B. S. Minhas and R. M. Solow. "Capital-Labor Substitution and Economic Efficiency," Review of Economics and Statistics, 43 (1961) :225-250.

Berndt, Ernst and Laurits Christensen. "The Translog Function and the Substitution of Equipment, Structures and Labor in U.S. Manufacturing 1929-68," Journal of Econometrics, Vol. 1 (1973) :81-114.

Berndt, E. R. and D. O. Wood. "Technology, Prices and the Derived Demand for Energy," Review of Economics and Statistics, 57 (1975) :259-68.

Binswanger, Hans P. "A Cost Function Approach to the Measurement of Elasticities of Substitution," American Journal of Agricultural Economics, 56 (1974) :377-86.

__________. "The Measurement of Technical Change Biases with Many Factors of Production," American Economic Review, 61 (1971) :94-109.

__________. "The Use of Duality Between Production, Profit and Cost Functions in Applied Econometric Research: A Didactic Note," Occastional Paper No. 10, Economics Department, 'ICRISAT', Hyderabad 500 016, India, July 1975.

Burgess, D. F. "Duality Theory and Pitfalls in the Specification of Technologies." Journal of Econometrics, 3 (1975) :105-21.

Christensen, L., D. Jorgenson, and L. Lau. "Conjugate Duality and Transcendental Logarithmic Production Frontiers," University of Wisconsin, July 1972.

__________. "Transcendental Logarithmic Production Frontiers," Review of Economics and Statistics, 54 (1972) :28-42.

Diewart, W. E. "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function," Journal of Political Economy, 79 (1971) :481-507.

Douglas, P. H. "Cobb-Douglas Production Function Once Again: Its History, Its Testing and Some New Empirical Values," Journal of Political Economy, 84 (1976) :903-15.

Dunne, James. "Estimating the Use Value of Agricultural Land for Farm Assessments in New York State." Unpublished Ph.D. thesis, Cornell University, 1981.

Dunne, J. and R. N. Boisvert. An Empirical Evaluation of Alternative Procedures for Determining Agricultural Use Values in New York, A.E. Res. 82-10, Department of Agricultural Economics, Cornell University, March 1982.

Ferguson, C. E. The Neoclassical Theory of Production and Distribution, Cambridge: Cambridge University Press, 1969.

Frisch, R. The Theory of Production, Chicago: Rand-McNally, 1965.

Goldman, S. M. and H. Uzawa. "A Note on Separability in Demand Analysis." Econometrica 32 (1964) :387-98.

Griliches, Z. and V. Ringstad. Economies of Scale and the Form of the Production Function, Amsterdam: North-Holland Publishing Co., 1971.

Hicks, J. "Elasticity of Substitution Again:: Substitutes and Complements," Oxford Economic Papers, 22 (1970) :289-96.

Humphrey, D. and J. Moroney. "Substitution among Capital, Labor and Natural Resource Products in American Manufacturing," Journal of Political Economy, 83 (1975) :57-82.

Kmenta, J. "On Estimation of the CES Production Function," International Economic Review, 8 (1967) :180-89.

Kmenta, J. and R. Gilbert. "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions," Journal of American Statistical Association, 63 (1968) 1180-2000.

Leontief, W. "Introduction to a Theory of the Internal Structure of Functional Relationships," Econometrica, 15 (1947) :361-73.

Maddala, G. S. Econometrics, New York: McGraw-Hill Book Co., 1977.

Miller, M., N. Whittlesey and T. Barr. "The Constant Elasticity of Substitution Production Function and Its Application in Research," Technical Bulletin Agricultural Research Center, Washington State University, August 1975.

Minhas, B. "The Homohypallagic Production Function, Factor-Intensity Reversals, and the Heckscher-Ohlin Theorem," Journal of Policital Economy, 70 (1962) :138-56.

Ruble, W. "Improving the Computation of Simultaneous Stochastic Linear Equation Estimates," Ph.D. dissertation, Michigan State University, 1968.

Sargan, J. D. "Production Functions: Summary, Theoretical Considerations and Results," in Layard et al., ed. Qualified Manpower and Economic Performance, London: Allan Love, 1971.

Sato, R. and T. Koizumi. "On the Elasticity of Substitution and Complementarity," Oxford Economic Papers, 25 (1973) :44-56.

Shih, J., L. Hushak and N. Rask. "The Validity of the Cobb-Douglas Specifications in Taiwan's Developing Agriculture." American Journal of Agricultural Economics, 59 (1977) :554-58.

Solow, R. "The Production Function and the Theory of Capital." Review of Economic Studies, 22 (1954-55) :101-08.

Theil, H. A System-wide Approach to Microeconomics. Chicago: The University of Chicago Press, 1980.

Uzawa, H. "Production Functions with Constant Elasticities of Substitution," Review of Economics Studies, 29 (1962) :291-99.

Vinod, H. "Nonhomogeneous Production Functions and Applications to Telecommunications," Bell Journal of Economics and Management Science, 3 (1972) :531-43.

Wyzan, M. "Empirical Analysis of Soviet Agricultural Production and Policy," American Journal of Agricultural Economics, 63 (1981) :475-483.

Zellner, A., J. Kmenta and J. Dreze. "Specification and Estimation of Cobb-Douglas Production Function Models," Econometrica, 34 (1966) :784-95.

APPENDIX A

In order to test the hypothesis that the translog production function is linear homogeneous, one must first distinguish between the algebraic form in equation (2) and the form in which it would be estimated. This is necessary because econometric methods do not distinguish between $\ln x_i \ln x_j$ and $\ln x_j \ln x_i$. Therefore, either β_{ij} or β_{ji} $(i \neq j)$ are estimated, but not both as is implied by equation (2). This presents no problem because $\beta_{ij} = \beta_{ji}$. The estimating equation is

$$\text{(1A)} \quad \ln y_t = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_{it} + \sum_{i=1}^{n} \gamma_{ii} (\ln x_{it})^2 + \sum_{i<j} \sum \beta_{ij} \ln x_{it} \ln x_{jt} + u_t$$

where

$t = 1, \ldots, T$, the number of observations; and

u_t = is a normally distributed error term with zero mean and finite variance

$\gamma_{ii} = 1/2\ \beta_{ii}$ from equation (2).

The hypothesis test can be stated in terms of the estimated parameters (and $n = 3$, the number of inputs)

H_0: $H\hat{\beta} = h$

H_1: $H\hat{\beta} \neq h$

where

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \ln \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \hat{\gamma}_{11} \\ \hat{\beta}_{12} \\ \hat{\beta}_{13} \\ \hat{\gamma}_{22} \\ \hat{\beta}_{23} \\ \hat{\gamma}_{33} \end{bmatrix} ; \text{ and } \quad h = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} .$$

Thus,

$$H\hat{\beta} = \begin{bmatrix} \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 \\ 2\hat{\gamma}_{11} + \hat{\beta}_{12} + \hat{\beta}_{13} \\ \hat{\beta}_{21} + 2\hat{\gamma}_{22} + \hat{\beta}_{23} \\ \hat{\beta}_{31} + \hat{\beta}_{32} + 2\hat{\gamma}_{33} \end{bmatrix}$$

The test statistic is

$$\text{(2A)} \quad \underset{(r,t-k)}{F} = \frac{\underset{(1xr)}{[H\hat{\beta} - h]} \left[\underset{(rxk)}{H} \ \underset{(kxk)}{(\underline{\overline{X}}'\underline{\overline{X}})^{-1}} \ \underset{(kxr)}{H'} \right]^{-1} \underset{(rx1)}{[H\hat{\beta} - h]/r}}{\frac{\hat{u}'\hat{u}}{T-k}}$$

where

$\underline{\overline{X}}$ = the matrix of observations on the independent variables in (1A), including the constant and all linear ($\ln x_i$), squared $(\ln x_i)^2$ and cross product ($\ln x_i \ln x_j$) terms;

$\hat{u}$ = vector of estimated residuals;

T = number of observations;

k = number of independent variables, including the constant (in the above example k = 10); and

r = number of linear restrictions (r = 4 in the above example).

An equivalent way to test for homogeneity is to estimate the model in (1A) in an unrestricted form, reestimate it imposing the restrictions implied in H_0 and compare the residual sums of squares for the restricted model (RRSS) with the residual sums of squares for the unrestricted model (URSS) (Maddala, 1977, p. 197). The test statistic is

$$(3A)\quad F_{(r,T-k)} = \frac{(RRSS-URSS)/r}{URSS/(T-k)} .$$

In order to perform this test, the restricted model can be estimated by beginning with the following algebraic form (for n=3)[1]

$$\begin{aligned}(4A)\quad \ln y = \ln \alpha_0 &+ \alpha_1 \ln x_{1t} + \alpha_2 \ln x_{2t} + \alpha_3 \ln x_{3t} + 1/2\beta_{11}(\ln x_{1t})^2 \\ &+ \beta_{12} \ln x_{1t} \ln x_{2t} + \beta_{13} \ln x_{1t} \ln x_{3t} + 1/2\beta_{22}(\ln x_{2t}^2) \\ &+ \beta_{23} \ln x_{2t} \ln x_{3t} + 1/2\beta_{33}(\ln x_{3t}^2) + \varepsilon_t\end{aligned}$$

The homogeneity and symmetry conditions require

$$(5A)\quad \beta_{11} + \beta_{12} + \beta_{13} = 0 \rightarrow \beta_{13} = -\beta_{11} - \beta_{12};$$

$$(6A)\quad \beta_{12} + \beta_{22} + \beta_{23} = 0 \rightarrow \beta_{23} = -\beta_{12} - \beta_{22};$$

$$(7A)\quad \beta_{13} + \beta_{23} + \beta_{33} = 0 \rightarrow \beta_{33} = -\beta_{13} - \beta_{23} = \beta_{11} + 2\beta_{12} + \beta_{22};$$

$$(8A)\quad \beta_{12} = \beta_{21};\ \beta_{13} = \beta_{31};\ \text{and}\ \beta_{23} = \beta_{32}$$

---

[1] The procedure for doing this was suggested by T. D. Mount.

Using equations (5A)-(8A), one can eliminate all the parameters in equation (2) except the α's and β_{11}, β_{12} and β_{22}.

$$\text{(9A)} \quad \ln y = \ln \alpha_0 + \alpha_1 \ln x_{1t} + \alpha_2 \ln x_{2t} + \alpha_3 \ln x_{3t} + 1/2\ \beta_{11}(\ln x_{1t})^2$$
$$+ \beta_{12} \ln x_{1t} \ln x_{2t} + (-\beta_{11}-\beta_{12})(\ln x_1 \ln x_{3t}) + 1/2\ \beta_{22}(\ln x_{2t})^2$$
$$+ (-\beta_{12}-\beta_{22})(\ln x_{2t} \ln x_{3t}) + 1/2(\beta_{11} + 2\beta_{12} + \beta_{22})(\ln x_{3t})^2$$
$$+ \varepsilon_t' .$$

Written another way, one obtains

$$\text{(10A)} \quad \ln y = \ln \alpha_0 + \alpha_1 \ln x_{1t} + \alpha_2 \ln x_{2t} + \alpha_3 \ln x_{3t}$$
$$+ \beta_{11}\ [1/2(\ln x_{1t})^2 - \ln x_{1t} \ln x_{3t} + 1/2(\ln x_{3t})^2]$$
$$+ \beta_{12}\ [\ln x_{1t} \ln x_{2t} - \ln x_{1t} \ln x_{3t} - \ln x_{2t} \ln x_{3t} + (\ln x_{3t})^2]$$
$$+ \beta_{22}\ [1/2(\ln x_{2t})^2 - \ln x_{2t} \ln x_{3t} + 1/2(\ln x_{3t})^2] + \varepsilon_t'.$$

Linear homogeneity can be tested by also placing restrictions on the sum of the α_i's. This model provides a way of estimating α_1, α_2, α_3, β_{11}, β_{12} and β_{13} by regressing ln y on against some transformed variables.

The values for the remaining restricted parameters can be obtained by substituting into (5A)-(8A). The restricted residual sums of squares can be used in the test described by (3A).

APPENDIX B

This appendix focuses on the mechanical problems associated with estimating the CES production function and the extent to which the translog formulation can be used to approximate the CES function.

In one of its most general forms, the CES function can be written as

$$(1B)\quad y = \alpha_0\left[\sum_{i=1}^{n} a_i\, x_i^{-\rho}\right]^{-v/\rho};\ \sum a_i = 1;\ a_i > 0;\ \alpha_0 > 0;\ \infty \geq \rho \geq -1;\ v > 0.$$

It can be shown that in this form, the scale economies are equal to v and the Allen partial elasticities of substitution (for v = 1) are $\sigma_{ij} = 1/1+\rho$ (Ferguson, 1969).

From the point of view of estimating a stochastic version of this function, it is neither linear in the parameters nor is it linear in logarithms.

$$(2B)\quad \ln y = \ln \alpha_0 - (v/\rho)\ \ln\left[\sum_{i=1}^{n} a_i\, x_i^{-\rho}\right].$$

Thus it cannot be estimated by ordinary least squares. Because economists throughout the 1960's were hopeful that the CES function would allow for a more meaningful set of empirical investigations, they searched for alternative ways to estimate directly or approximate the parameters of such a production function.

The direct estimation of the parameters involves the use of non-linear least-squares techniques relying on one of two approaches. The model is either:

a) expanded as a Taylor series, with corrections made to the several parameters at successive iterations, or

b) methods of steepest descent are used to minimize the sum's of squared deviations.

The Taylor Approximations often diverge and the gradient methods converge slowly. According to Miller et al. (1975), Marquardt's algorithm, based on the maximum neighborhood method, combines the best of both of the above. However, neither Miller et al. nor others have had much success. There is difficulty in selecting initial values for the parameters and the methods still do not converge. Furthermore, these estimates are expensive to obtain and are still only approximations.

If direct estimation does not work, what are the alternatives? One of them was proposed in part by Arrow et al. (1961) in their original article. For the special case where v = 1 and n = 2 from equation (1A) we have by linear homogeneity and the assumption of competitive equilibrium

(3B) $\ln \frac{y_t}{x_{2i}} = \ln \alpha + \beta \ln P_{2t} + \varepsilon_t$ (where P_{2i} is the normalized price of x_2), where they demonstrate that the direct elasticity of substitution between the two inputs σ is given by

(4B) $\sigma = \frac{1}{1+\rho} = \beta;$

thus, by estimating β using ordinary least squares, one could find

(5B) $\hat{\rho} = [\frac{1}{\hat{\beta}}] - 1.$

Having identified $\hat{\rho}$ one might be tempted to try and estimate a_1 and a_2 by plugging (5B) into a stochastic version of equation (2B) for n = 2 and v = 1:

6B) $\ln y_t = \ln \alpha_0 + (-1/\hat{\rho}) \ln [a_1 x_{1t}^{-\hat{\rho}} + a_2 x_{2t}^{-\hat{\rho}}] + e_t.$

But even if we have an independent estimate of ρ, one still cannot identify the parameters of a_1 and a_2 in this second step. One is left with no way to estimate the distribution parameters. Furthermore, even the parameter ρ cannot be estimated in this fashion for $v \neq 1$.

Kmenta (1967) and Griliches and Ringstad (1971) have suggested an alternative based on a particular second order Taylor series approximation. The approximation is about $\rho = 0$.

To illustrate, use the two variable case. In Kmenta's notation, y is a function of K and L.

(7B) $\ln y = \ln \gamma - v/\rho \ln[\delta K^{-\rho} + (1-\delta)L^{-\rho}]$.

Because one wants to expand equation (7B) about $\rho = 0$, consider it as a function of ρ. Letting

(8B) $f(\rho) = \ln(\delta K^{-\rho} + (1-\delta)L^{-\rho})$.

so that

(9B) $\ln y = \ln \gamma - v/\rho\ [f(\rho)]$.

All one has to do is find an approximation for $f(\rho)$ expanded around $\rho = 0$. Evaluate

(10B) $f(o) = \ln[\delta K^{-o} + (1-\delta)L^{-o}] = \ln 1 = 0$; if $0 < \delta < 1$

and using the laws of differentiation[1]

(11B) $f'(\rho) = \dfrac{1}{[\delta K^{-\rho} + (1-\delta)L^{-\rho}]} [-\delta K^{-\rho} \ln K - (1-\delta)^{-\rho} \ln L]$

and

(12B) $f'(o) = -[\delta \ln K + (1-\delta)\ln L]$.

We also have

(13B) $f''(\rho) = [\delta K^{-\rho} + (1-\delta)L^{-\rho}]^{-1} [\delta K^{-\rho} (\ln K)^2 +$

$(1-\delta)L^{-\rho} (\ln L)^2] - [-\delta K^{-\rho} \ln K - (1-\delta)L^{-\rho} \ln L]$

$[\delta K^{-\rho} + (1-\delta)L^{-\rho}]^{-2} [-\delta K^{-\rho} \ln K - (1-\delta)L^{-\rho} \ln L]$.

---

[1]If u is a function of x and c is a constant, then

$$\frac{dc^u}{dx} = c^u \frac{du}{dx} \ln c.$$

Evaluated at $\rho = 0$,

$$(14B)\ f''(o) = [1]\ [\delta(\ln K)^2 + (1-\delta)(\ln L)^2]$$
$$+ [\delta \ln K + (1-\delta)\ln L][1][-\delta \ln K - (1-\delta)(\ln L)]$$
$$= \delta(\ln K)^2 + (1-\delta)(\ln L)^2 + [-\delta^2(\ln K)^2 - (1-\delta)\delta \ln K \ln L$$
$$- \delta(1-\delta)\ln K \ln L - (1-\delta)^2(\ln L)^2]$$
$$= \delta(1-\delta)(\ln K)^2 - 2\delta(1-\delta)(\ln K)(\ln L)$$
$$+ (1-\delta)(\ln L)^2 - (1-2\delta+\delta^2)(\ln L)^2$$
$$= \delta(1-\delta)[(\ln K)^2 - 2 \ln K \ln L + (\ln L)^2]$$
$$= \delta(1-\delta)[\ln K - \ln L]^2.$$

Substitution into the Taylor's Approximation gives[2]

$$(15B)\ f(\rho) \cong -\rho(\delta \ln K + (1-\delta)\ln L) + 1/2\rho^2\ \delta(1-\delta)[\ln K - \ln L]^2$$

and

$$(16B)\ \ln y \cong \ln \gamma + v(\delta \ln K + (1-\delta)\ln L) - 1/2v\rho\ \delta(1-\delta)[\ln K - \ln L]^2.$$

The extension of this idea to the function in n variables[3]

---

[2]The expansion around $\rho = 0$ is

$$f(\rho) = f(o) + f'(o)(\rho-o) + \frac{f''(o)}{2}(\rho-o)^2 .$$

[3]To derive the second order Taylor's approximation to the general CES formulation in (2B) write it in logarithms as

$$(1')\quad \ln y = \ln \alpha_0 - v/\rho\ g(\rho)$$

where $g(\rho) = \ln [\sum_{i=1}^{n} a_i x_i^{-\rho}]$.

We need to evaluate

$$(2')\quad g(o) = \ln [\sum_{i=1}^{n} a_i x_i^{-0}] = \ln [\sum_{i=1}^{n} a_i] = \ln 1 = 0, \text{ since } \sum_{i=1}^{n} a_i = 1.$$

(17B) $y = \alpha_0 \left[\sum_{i=1}^{n} a_i x_i^{-\rho} \right]^{-v/\rho}; \sum a_i = 1$ is given by

(18B) $\ln y = \ln \alpha_0 + v \sum_{i=1}^{n} a_i \ln x_i - \frac{v\rho}{2} \left[\sum a_i (\ln x_i)^2 - (\sum a_i \ln x_i)^2 \right]$

or in a more recognizable form

(footnote (3) cont.)

We also need

$$(3') \quad g'(\rho) = \left[\frac{1}{\sum_{i=1}^{n} a_i x_i^{-\rho}} \right] \left[-\sum_{i=1}^{n} a_i x_i^{-\rho} \ln x_i \right]$$

$$(4') \quad g'(0) = \left[\frac{1}{\sum_{i=1}^{n} a_i} \right] \left[-\sum_{i=1}^{n} a_i \ln x_i \right] = \left[-\sum_{i=1}^{n} a_i \ln x_i \right].$$

Using (3') we can find

$$(5') \quad g''(\rho) = \left[\sum_{i=1}^{n} a_i x_i^{-\rho} \right] \left[\sum_{i=1}^{n} a_i x_i^{-\rho} (\ln x_i)^2 \right] + \left[-\sum_{i=1}^{n} a_i x_i^{-\rho} \ln x_i \right] \left[-\sum_{i=1}^{n} a_i x_i^{-\rho} \right]^{-2} \left[-\sum_{i=1}^{n} a_i x_i^{-\rho} \ln x_i \right]$$

$$(6') \quad g''(0) = (1) \left[\sum_{i=1}^{n} a_i (\ln x_i)^2 \right] - \left[\sum_{i=1}^{n} a_i \ln x_i \right]^2.$$

Using equations (2'), (4') and (6'), we have the second order approximation

$$g(\rho) \cong -\rho \left[\sum_{i=1}^{n} a_i \ln x_i \right] + \frac{\rho^2}{2} \left[\sum_{i=1}^{n} a_i (\ln x_i)^2 - \left(\sum_{i=1}^{n} a_i \ln x_i \right)^2 \right].$$

Plugging this into (1') we have

$$(7') \quad \ln y = \ln \alpha_0 + v \left[\sum_{i=1}^{n} a_i \ln x_i \right] - \frac{v\rho}{2} \left[\sum_{i=1}^{n} a_i (\ln x_i)^2 - \left(\sum_{i=1}^{n} a_i \ln x_i \right)^2 \right].$$

This is the same as (18B).

$$\text{(19B)} \quad \ln y = \ln \alpha_0 + v \sum_{i=1}^{n} a_i \ln x_i - \frac{v\rho}{2} [\sum_{i=1}^{n} a_i(1-a_i)(\ln x_i)^2 + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} a_i a_j \ln x_i \ln x_j].$$

By letting $a_{ii} = a_i(1-a_i)\frac{\rho v}{2}$ and $a_{ij} = \frac{v\rho}{2} a_i a_j$,

we know that $a_{ij} = a_{ji}$ and

$$\text{(20B)} \quad \ln y = \ln \alpha_0 + v \sum_{i=1}^{n} a_i \ln x_i - [\sum_{i=1}^{n} a_{ii} \ln x_i^2 + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} a_{ij} \ln x_i \ln x_j$$

Both equations (16B) and (20B) are translog functions and if one assumes stochastic forms of the equations, then the composite parameters can be estimated. There are two perspectives from which we can view this approximation and the estimated parameters. First, they provide a test of the Cobb-Douglas Form. Second, they are potentially a way to estimate the parameters of a CES production function.

To begin, the approximation (which is a quadratic function in logarithms) is better the closer σ is to unity because the Taylor's expansion is about $\rho = 0$. This is not surprising since the approximations (equations (16B) or (18B)) reduce to the Cobb-Douglas function when $\rho = 0$. Thus, even though the quality of the approximation deteriorates as σ departs from unity, a standard F-test provides a direct test of the Cobb-Douglas form (e.g., test the terms on the quadratic expressions).

The F-test is one involving the C-D equation (the translog model with the coefficients on the non-linear terms restricted to zero) and the unrestricted model.

H_0: $a_{ii} = a_{ij} = 0$; all i,j in equation (20B)

H_1: all else

The test is

$$(21B)\quad F_{r,(n-m-1)} = \frac{(RRSS-URSS)/r}{(URSS)/n-m-1}$$

where

RRSS = restricted residual sums of squares

URSS = unrestricted residual sums of squares

r = number of restrictions

n = observations

m = independent variables.

There are two minor problems, however, in interpreting this test. Because the test is based only on a second order approximation to the CES, one has no way to know the effects of these omitted terms. Thus, it is not all clear that a rejection of the C-D hypothesis is necessarily confirmation of the CES hypothesis. One might, for example, argue that the alternative hypothesis is not restricted to the CES form. The H_1 is consistent with the hypothesis of a general but unspecified form.

The second problem is that the distribution parameters (a_i's) are embodied in these terms. These are $0 \leq a_i \leq 1$, the absolute value of the complex parameters is likely to be small. Thus, a large sample and good variation in the input levels are needed for any confidence in the estimates.

More importantly, there are major difficulties in identifying the underlying CES parameters from the composite parameters of (16B) and (20B). For the two input model, Kmenta (1967) approaches the problem in several different ways. First, in estimating equation (16B), Kmenta follows the lead of Arrow et al., predetermines $\delta = 0.519$ and applies simple least squares. He says this was to avoid multicollinearity in time series data, but in general, there seems little basis on which to set δ *a priori*.

The second approach involves estimating the parameters of (7B) from the parameters of (16B). One knows that following:

(22B) $\widehat{\ln} \gamma$

from (16B) is a direct estimate of $\ln \gamma$ in equation (7B).

The parameters of (7B) in terms of the estimated parameters of (16B) are:

(23B) $v\delta = \hat{\beta}_1$

(24B) $v(1-\delta) = \hat{\beta}_2$

(25B) $\rho \; v \; \delta(1-\delta) = -2\hat{\beta}_3.$

To see if they are identified (eliminating the "hats" for simplicity)

(26B) $(v-v\delta) = \beta_2$

plugging in from (23B), one has

(27B) $v - \beta_1 = \beta_2$; and

(28B) $v = \beta_1 + \beta_2$; and $\delta = \dfrac{\beta_1}{\beta_1 + \beta_2}$

(In (25B) and from (23B)

(29B) $\rho(\beta_1 + \beta_2)(\dfrac{\beta_1}{\beta_1 + \beta_2}) \; (1 - \dfrac{\beta_1}{\beta_1 + \beta_2}) = -2\beta_3;$

(30B) $\rho(\beta_1)\,(1 - \frac{\beta_1}{\beta_1 + \beta_2}) = -2\beta_3;$

(31B) $\rho(\frac{\beta_1(\beta_1 + \beta_2) - \beta_1^2}{\beta_1 + \beta_2}) = -2\beta_3;$

(32B) $\rho = -2[\frac{\beta_1 + \beta_2}{\beta_1\beta_2}]\beta_3;$

In this case, we can, without predetermining the value of δ, identify ρ, v, δ from (28B), (23B) and (32B).

Kmenta (1967) also examines two simultaneous equation estimates for the parameters of the two input model. Since these are readily available in the literature and present few problems, the discussion now focuses on the general case, equation (20B) in this case. The following parameters are estimated

(33B) $\hat{a}_i$, $\hat{a}_{ii}$ for (i=1,...,n)

(34B) $\hat{a}_{ij}$ for $i \neq j$.

There are $2n + \frac{n(n-1)}{2}$ original parameters to be estimated (equation (17B)). Direct estimates of the a_i's are obtained but then

(35B) $a_{ii} = a_i(1-a_i)\,\frac{\rho v}{2}$; (i=1,...,n)

(36B) $a_{ij} = \frac{v\rho}{2}\,a_i\,a_j = a_{ji}$ for all i,j; and

(37B) $\sum_{i=1}^{n} a_i = 1.$

Clearly v and ρ are overidentified and there is no straightforward way of placing a restriction on the sum of the a_i's. Thus, the translog form is of limited use in estimating the parameters of the CES production function, except in the two input case.

APPENDIX C

To discuss the elasticity of substitution between two inputs in the translog case, this appendix focuses on the direct elasticity of substitution as defined by Sargan (1971) or Sato and Koizumi (1973). Within this context, one can demonstrate that the direct elasticity of substitution for equation (2) is a tremendously complex function of the parameters and the input levels. During the derivation, it is shown that if the variables are scaled around a point before the parameters are estimated, then the direct elasticity of substitution <u>at that point</u> is a function of the parameters only. The practical implication of this fact is that if one is interested only in the characteristics of the production function at several points, it may be easier to re-estimate the scaled function at these points and calculate the appropriate attributes of the function then it is to work with the unscaled function. To begin, let's reiterate equation (2):

$$(1C)\quad \ln y = \ln \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + 1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j .$$

The formula for the "direct elasticity of substitution" is given by

$$(2C)\quad e_{ij} = \frac{d(x_j/x_i)}{x_j/x_i} \div \frac{d(f_i/f_j)}{f_i/f_j} .$$

Letting

$$(3C)\quad t = \frac{\partial y/\partial x_i}{\partial y/\partial x_j} = \frac{f_i}{f_j} = \frac{x_j(\alpha_i + \sum_k \beta_{ik} \ln x_k)}{x_i(\alpha_j + \sum_k \beta_{jk} \ln x_k)} = -\frac{dx_j}{dx_i}$$

$$= (\alpha_i x_j + x_j \sum_k \beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum_k \beta_{jk} \ln x_k)^{-1}$$

then,

$$(4C)\quad dt = \frac{\partial t}{\partial x_j}\,dx_j + \frac{\partial t}{\partial x_i}\,dx_i$$

Therefore,[1]

$$(5C)\quad \frac{\partial t}{\partial x_j} = (\alpha_j x_i + x_i \textstyle\sum \beta_{jk} \ln x_k)^{-1}(\alpha_i + \sum \beta_{ik} \ln x_k + \beta_{ij})$$
$$- (\alpha_i x_j + x_j \textstyle\sum \beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum \beta_{ik} \ln x_k)^{-2}(x_i)(\beta_{jj}/x_j)$$

and

$$(6C)\quad \frac{\partial t}{\partial x_i} = (\alpha_j x_i + x_i \textstyle\sum \beta_{jk} \ln x_k)^{-1}(x_j)(\beta_{ii}/x_i)$$
$$- (\alpha_i x_j + x_j \textstyle\sum \beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum \beta_{jk} \ln x_k)^{-2}$$
$$(\alpha_j + \textstyle\sum \beta_{jk} \ln x_k + \beta_{ji}).$$

From equations (3C) and (4C), we have the equations in table 1C, equations (7C) and (8C). If one defines the production elasticity as

$$(9C)\quad e_i = \alpha_i + \textstyle\sum \beta_{ik} \ln x_k,$$

one obtains equation (10C) in table 1C by simple substitution.

To complete the derivation of e_{ij}, one must also find

$$(11C)\quad d(x_j/x_i) = (1/x_i)dx_j - (x_j/x_i^2)dx_i$$

and

$$(12C)\quad \frac{d(x_j/x_i)}{x_j/x_i} = (1/x_j)dx_j - (1/x_i)dx_i.$$

Substituting (12C) and (10C) into (2C), one obtains equation (13C) in table 2C. In order to be useful, however, we must eliminate dx_i and dx_j from equation (13C). We can do this by recalling from (3C) that

$$(14C)\quad -\frac{dx_j}{dx_i} = \frac{x_j e_i}{x_i e_j}$$

(which guarantees that output is constant) and

$$(15C)\quad dx_j = -\frac{x_j e_i}{x_i e_j}\,dx_i.$$

---

[1]Unless otherwise specified, the index of summation is k throughout this appendix.

Table 1C: Equations for dt/t

$$(7C)\quad \frac{dt}{t} = \frac{[(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-1}(\alpha_i + \sum\beta_{ik} \ln x_k + \beta_{ij}) - (\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-2} \frac{x_i\beta_{jj}}{x_j}]dx_j}{(\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-1}}$$

$$+ \frac{[(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-1} \frac{x_j\beta_{ii}}{x_i} - (\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-2}(\alpha_j + \sum\beta_{jk} \ln x_k + \beta_{ji})]dx_i}{(\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-1}}$$

$$(8C)\quad \frac{dt}{t} = \frac{\{[(\alpha_i + \sum\beta_{ik} \ln x_k + \beta_{ij}) - (\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-1} \frac{x_i\beta_{jj}}{x_j}]dx_j\}}{(\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)}$$

$$+ \frac{\{\frac{x_j\beta_{ii}}{x_i} - (\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)(\alpha_j x_i + x_i \sum\beta_{jk} \ln x_k)^{-1}(\alpha_j + \sum\beta_{jk} \ln x_k + \beta_{ji}]dx_i\}}{(\alpha_i x_j + x_j \sum\beta_{ik} \ln x_k)}$$

$$(10C)\quad \frac{dt}{t} = \frac{\{[(e_i + \beta_{ij}) - (x_j e_i)(x_i e_j)^{-1}(\frac{x_i\beta_{jj}}{x_j})]dx_j\}}{x_j e_i} + \frac{\{[(\frac{x_j\beta_{ii}}{x_i}) - (x_j e_i)(x_i e_j)^{-1}(e_j + \beta_{ji})]dx_i\}}{x_j e_i}$$

Table 2C. Equations for e_{ij}.

(13C)
$$e_{ij} = \frac{x_j e_i\{(1/x_j)dx_j - (1/x_i)dx_i\}}{\{[(e_i + \beta_{ij}) - (x_j e_i)(x_i e_j)^{-1}(\frac{x_i \beta_{jj}}{x_j})]dx_j\} + \{[(\frac{x_j \beta_{ii}}{x_i}) - (x_j e_i)(x_i e_j)^{-1}(e_j + \beta_{ji})]dx_i\}}$$

(16C)
$$e_{ij} = \frac{x_j e_i\{-\frac{1}{x_j}[\frac{x_j e_i}{x_i e_j}]dx_i - (\frac{1}{x_i})dx_i\}}{\{[(e_i + \beta_{ij}) - (x_j e_i)(x_i e_j)^{-1}(\frac{x_i \beta_{jj}}{x_j})][-\frac{x_j e_i}{x_i e_j}]dx_i\} + \{[(\frac{x_j \beta_{ii}}{x_i}) - (x_j e_i)(x_i e_j)^{-1}(e_j + \beta_{ji})]dx_i\}}$$

(17C)
$$e_{ij} = \frac{a_i\{-\frac{a_i}{a_j} - 1\}}{\{(a_i + b_{ij} - \frac{a_i}{a_j}b_{jj})(\frac{-a_i}{a_j})\} + \{(b_{ii} - (\frac{a_i}{a_j})(a_j + b_{ji})\}} = \frac{-\frac{a_i}{a_j}\{a_i + a_j\}}{-[\frac{a_i^2}{a_j} + \frac{b_{ij}a_i}{a_j} - \frac{a_i^2 b_{jj}}{a_j^2}] + [b_{ii} - a_i - \frac{a_i b_{ij}}{a_j}]}$$

$$= \frac{-\frac{a_i}{a_j}\{a_i + a_j\}}{\frac{-2b_{ij}a_i}{a_j} - \frac{a_i^2}{a_j} + \frac{a_i^2 b_{jj}}{a_j^2} + b_{ii} - a_i} = \frac{-\frac{1}{a_j}\{a_i + a_j\}}{\frac{-2b_{ij}a_i}{a_j a_i} - \frac{a_i^2}{a_j a_i} + \frac{a_i^2 b_{jj}}{a_j^2 a_i} + \frac{b_{ii}}{a_i} - \frac{a_i}{a_i}}$$

$$= \frac{-\{a_i + a_j\}}{\frac{-2b_{ij}a_i a_j}{a_i a_j} - \frac{a_i^2 a_j}{a_i a_j} + \frac{a_i^2 b_{jj}}{a_j a_i} + \frac{a_j^2 b_{ii}}{a_i a_j} - \frac{a_i a_j^2}{a_i a_j}} = \frac{-\{a_i + a_j\}}{-(a_i + a_j) + [\frac{a_i^2 b_{jj} - 2b_{ij}a_i a_j + b_{ii}a_j^2}{a_i a_j}]}$$

(Note: This is the negative of Sargan's (1971) definition. Under this development, a function with convex isoquants has $e_{ij} > 0$.)

Substituting (15C) into (13C) we obtain (16C). This expression is the direct elasticity of substitution between x_i and x_j and for the model in equation (1C) it is a function of the α_i's, β_{ij}'s, and the input levels (the dx_i's cancel). It is not constant for all input levels. However, recall that in the case of interpretation 2, the scaled model (equation (17)) one knows that (from equation (25))

$$a_i = e_i \quad \text{for } (i=1,\ldots,n) \text{ and } b_{ij} = \beta_{ij}; \quad \text{all } i \text{ and } j$$

when e_i is evaluated at the geometric mean. Therefore, in order to estimate e_{ij}, one could estimate a scaled function and use equation (17C). In the scaled model also, $x_i = 1$, all i, at the point where e_{ij} is evaluated. Because it was shown in the text that b_{ij}'s are invariant with respect to the point of scaling but that the a_i's are not, e_{ij} remains a function of the input levels.