Duality in Production

W. Erwin Diewert.[1]
Discussion Paper 18-02,
Vancouver School of Economics,
University of British Columbia,
Vancouver, B.C., Canada, V6T 1L4.
Email: erwin.diewert@ubc.ca

February 5, 2018.

Abstract

The paper reviews the application of duality theory in production theory. Duality theory turns out to be a useful tool for two reasons: (i) it leads to relatively easy characterizations of the properties of systems of producer derived demand functions for inputs and producer supply functions for outputs and (ii) it facilitates the generation of flexible functional forms for producer demand and supply functions that can be estimated using econometrics. The paper focuses on describing the properties of five functional forms that have been used in the production literature: (i) the Constant Elasticity of Substitution (CES), (ii) the Generalized Leontief, (iii) the translog, (iv) the Normalized Quadratic and (v) the Konüs Byushgens Fisher functional forms. The applications of GDP functions and joint cost functions to various areas of applied economics is explained.

Keywords

Production theory, duality theory, cost functions, production functions, joint cost functions, national product functions, GDP functions, variable profit functions, properties of producer demand and supply functions, Shephard's Lemma, Hotelling's Lemma, Samuelson's Lemma, flexible functional forms, estimation of technical progress, the valuation of public sector outputs, modeling monopolistic behavior, sunk costs.

JEL Classification Numbers

C02, C32, C43, D24, D42, D92, E01, E22, F11, H44, L51, M40, O47

[1] W. Erwin Diewert: School of Economics, University of British Columbia, Vancouver B.C., Canada, V6T 1Z1 and the School of Economics, UNSW Australia, Sydney 2052. Email: erwin.diewert@ubc.ca. The author thanks Robert Cairns, Kevin Fox, John Hartwick, Robert Inklaar, Peter Neary, Subhash Ray, Stephen Redding, Philip Vermeulen and Valentin Zelenyuk for helpful comments and the SSHRC of Canada for financial support.

1. Introduction

Duality theory is a very useful tool for estimating production functions or more generally, for estimating production possibilities sets. It is also useful in allowing one to derive the theoretical properties that differentiable derived producer demand for input and supply of output functions must satisfy if the producer is maximizing profits or minimizing costs. This chapter will illustrate these advantages of duality theory in the producer context.

Sections 2-6 below will focus on the case of one output, N input technologies. The multiple output and multiple input case will be considered in Sections 15-22.[2]

The one output, many input cost function is defined in Section 2 and in Section 3, the conditions on the production function that allow the cost function to completely describe the underlying technology are listed: this establishes the Shephard (1953) Duality Theorem between cost and production functions. Section 4 explains Shephard's Lemma; i.e., it shows why differentiating a cost function with respect to input prices generates the vector of cost minimizing input demand functions. If the cost function is twice continuously differentiable with respect to input prices, then Section 5 derives the properties that the system of cost minimizing input demand functions must satisfy. Section 6 looks at the duality between cost and production functions if production is subject to constant returns to scale; i.e., if the production function is homogeneous of degree one in inputs.

Sections 7-11 look at specific functional forms for the cost function. The five functional forms that are studied are (i) the Constant Elasticity of Substitution (CES), (ii) the Generalized Leontief, (iii) the Translog, (iv) the Normalized Quadratic and (v) the Konüs Byushgens Fisher (KBF) functional forms. The last four functional forms are *flexible functional forms*; i.e., they can provide a second order approximation to an arbitrary twice continuously differentiable unit cost function at any arbitrary price point.[3] A major problem with flexible functional forms is the *curvature problem*; i.e., an estimated flexible functional form for a unit cost function may violate the concavity in prices property that cost functions must satisfy. It turns out that the Normalized Quadratic and KBF functional forms are such that the correct curvature conditions can be imposed without destroying the flexibility of the functional form.[4]

---

[2] In Sections 2-21, it will be assumed that the producer takes prices as given constants in each period. Section 22 extends the analysis to the case of monopolistic behavior.

[3] Diewert and Wales (1993; 89-92) discuss some additional flexible functional forms that are not discussed here. These alternative functional forms have various problems.

[4] On a personal note, I did my thesis on flexible functional forms and, with the help of Daniel McFadden (my thesis advisor), I came up with the Generalized Leontief cost function as my first attempt at finding a "perfect" functional form that was flexible, parsimonious (i.e., had the minimal number of parameters to be estimated that would enable it to be flexible) and generated derived demand (or supply) functions that were either linear or close to linear in the unknown parameters in order to facilitate econometric estimation. I was a graduate student at Berkeley at the time (1964-1968) and I met frequently with Dale Jorgenson. He and his student at the time, Lawrence Lau, realized that instead of taking a quadratic form in the square roots of input prices, one could take a quadratic form in the logarithms of prices as a functional form for the logarithm of the cost function and the translog functional form was born. However, empirical applications

Section 12 introduces the concept of a *semiflexible functional form*. A major problem with the use of a flexible functional form is that it requires the estimation of roughly $N^2/2$ parameters if there are N inputs. The semiflexible concept reduces this large number of parameters in a sensible way.

Section 13 shows how piece-wise linear functions of time can be used to model technical progress in a more general manner than just using linear time trends in the demand functions. Section 14 shows how a flexible functional form can be generalized to achieve the second order approximation property at two sample points if we are estimating production functions in the time series context.[5]

Section 15 introduces Samuelson's (1953) *National Product Function* or the *variable profit function*. This function conditions on a vector of fixed inputs and maximizes the value of outputs less variable inputs. The comparative statics properties of this function are developed in Section 16. Sections 17-19 look at three flexible functional forms for this function: (i) the translog, (ii) the Normalized Quadratic and (iii) the KBF variable profit functions. The systems of estimating equations that these functional forms generate are also exhibited.

Sections 20-22 develop the properties of *joint cost functions*; i.e., these functions generalize the one output cost function to a cost function for multiple output producers. Section 21 looks at three flexible functional forms for this function: (i) the translog, (ii) the Normalized Quadratic and (iii) the KBF joint cost functions. The latter two functions have the property that the correct curvature conditions can be imposed on them without destroying their flexibility properties. Section 22 looks at applications of joint cost functions to: (i) problems associated with the measurement of the outputs of public sector producers in the System of National Accounts, (ii) the measurement of the efficiency of regulated utilities and (iii) the estimation of technology sets when producers have some monopoly power.

Section 23 concludes with a listing of three problems that are not addressed in this chapter and require further research.

---

of these functional forms soon showed that these functional forms had a drawback: it was not possible to impose the correct concavity or convexity properties on these flexible functional forms without destroying the flexibility of the functional form. In the 1980s, Diewert and Wales came up with the normalized quadratic functional form which was flexible, parsimonious and had the property that the correct curvature conditions could be imposed without impairing the flexibility property. However, in order to preserve the parsimony property, one had to pick a more or less arbitrary alpha vector and imbed it into the functional form as we will see later in this chapter. But different choices of alpha could generate perhaps substantially different estimates for demand and supply elasticities. The last flexible functional form that we will discuss in this chapter, the KBF functional form, overcomes this difficulty and hence completes our quest for the "perfect" flexible functional form.

[5] It should be noted that our analysis is geared to the time series context. Much of our analysis can be translated to the cross sectional context.

It may be useful to use this chapter as part of a course in microeconomic theory or in production theory. To facilitate this use, the author has added many straightforward problems that the instructor can assign to students. These problems are also an efficient way of extending the results presented in the main text.

2. Cost Functions: The One Output Case

The *production function* and the corresponding *cost* function play a central role in many economic applications. In the following section, we will show that under certain conditions, the cost function is a sufficient statistic for the corresponding production function; i.e., if we know the cost function of a producer, then this cost function can be used to generate the underlying production function.

Let the producer's *production function* f(x) denote the maximum amount of output that can be produced in a given time period, given that the producer has access to the nonnegative vector of inputs, $x \equiv [x_1,\ldots,x_N] \geq 0_N$.[6] If the production function satisfies the minimal regularity condition of continuity from above,[7] then given any positive output level y that the technology can produce and any strictly positive vector of input prices $p \equiv [p_1,\ldots,p_N] >> 0_N$, we can calculate the producer's *cost function* C(y,p) as the solution value to the following constrained minimization problem:

(1) $C(y,p) \equiv \min_x \{p^T x : f(x) \geq y \; ; \; x \geq 0_N\}$.

It turns out that the cost function C will satisfy the following 7 properties, provided that the production function is continuous from above:[8]

Theorem 1; Diewert (1993; 107-114)[9]: Suppose f is continuous from above. Then C defined by (1) has the following properties:

Property 1: C(y,p) is a *nonnegative* function.
Property 2: C(y,p) is *positively linearly homogeneous in p* for each fixed y; i.e.,

(2) $C(y,\lambda p) = \lambda C(y,p)$ for all $\lambda > 0$, $p >> 0_N$ and $y \in$ Range f (i.e., y is an output level that is producible by the production function f).

Property 3: C(y,p) is *nondecreasing in p* for each fixed $y \in$ Range f; i.e.,

---

[6] Notation: $x \geq 0_N$ means each component of the vector x is nonnegative, $x > 0_N$ means $x \geq 0_N$ and $x \neq 0_N$ and $x >> 0_N$ means each component of x is positive. $p^T x \equiv \Sigma_{n=1}^N p_n x_n$. Vectors are understood to be column vectors when it matters.
[7] We require that f be *continuous from above* for the minimum to the cost minimization problem to exist; i.e., for every output level y that can be produced by the technology (so that $y \in$ Range f), we require that the set of x's that can produce at least output level y (this is the upper level set $L(y) \equiv \{x : f(x) \geq y\}$) is a closed set in R^N.
[8] Note that this minimal regularity condition cannot be contradicted using a finite data set.
[9] For the history of closely related results, see Diewert (1974a; 116-120).

(3) $y \in$ Range f, $0_N << p^1 < p^2$ implies $C(y,p^1) \le C(y,p^2)$.

Property 4: C(y,p) is a *concave function of p* for each fixed $y \in$ Range f; i.e.,

(4) $y \in$ Range f, $p^1 >> 0_N$; $p^2 >> 0_N$; $0 < \lambda < 1$ implies

$$C(y,\lambda p^1+(1-\lambda)p^2) \ge \lambda C(y,p^1) + (1-\lambda)C(y,p^2).$$

Property 5: C(y,p) is a *continuous function of p* for each fixed $y \in$ Range f.
Property 6: C(y,p) is *nondecreasing in y* for fixed p; i.e.,

(5) $p >> 0_N$, $y^1 \in$ Range f, $y^2 \in$ Range f, $y^1 < y^2$ implies $C(y^1,p) \le C(y^2,p)$.

Property 7: For every $p >> 0_N$, C(y,p) is *continuous from below in y*; i.e.,

(6) $y^* \in$ Range f , $y^n \in$ Range f for n = 1,2,…, $y^n \le y^{n+1}$, $\lim_{n\to\infty} y^n = y^*$ implies

$$\lim_{n\to\infty} C(y^n,p) = C(y^*,p).$$

Proof of Property 1: Let $y \in$ Range f and $p >> 0_N$. Then

$$\begin{aligned} C(y,p) &\equiv \min_x \{p^Tx : f(x) \ge y \,;\, x \ge 0_N\} \\ &= p^Tx^* && \text{where } x^* \ge 0_N \text{ and } f(x^*) \ge y \\ &\ge 0 && \text{since } p >> 0_N \text{ and } x^* \ge 0_N. \end{aligned}$$

Proof of Property 2: Let $y \in$ Range f, $p >> 0_N$ and $\lambda > 0$. Then

$$\begin{aligned} C(y,\lambda p) &\equiv \min_x \{\lambda p^Tx : f(x) \ge y \,;\, x \ge 0_N\} \\ &= \lambda \min_x \{p^Tx : f(x) \ge y \,;\, x \ge 0_N\} && \text{since } \lambda > 0 \\ &= \lambda C(y,p) && \text{using the definition of } C(y,p). \end{aligned}$$

Proof of Property 3: Let $y \in$ Range f, $0_N << p^1 < p^2$. Then

$$\begin{aligned} C(y,p^2) &\equiv \min_x \{p^{2T}x : f(x) \ge y \,;\, x \ge 0_N\} \\ &= p^{2T}x^* && \text{where } f(x^*) \ge y \text{ and } x^* \ge 0_N \\ &\ge p^{1T}x^* && \text{since } x^* \ge 0_N \text{ and } p^2 > p^1 \\ &\ge \min_x \{p^{1T}x : f(x) \ge y \,;\, x \ge 0_N\} && \text{since } x^* \text{ is feasible for this problem} \\ &\equiv C(y,p^1). \end{aligned}$$

Proof of Property 4: Let $y \in$ Range f, $p^1 >> 0_N$; $p^2 >> 0_N$; $0 < \lambda < 1$. Then

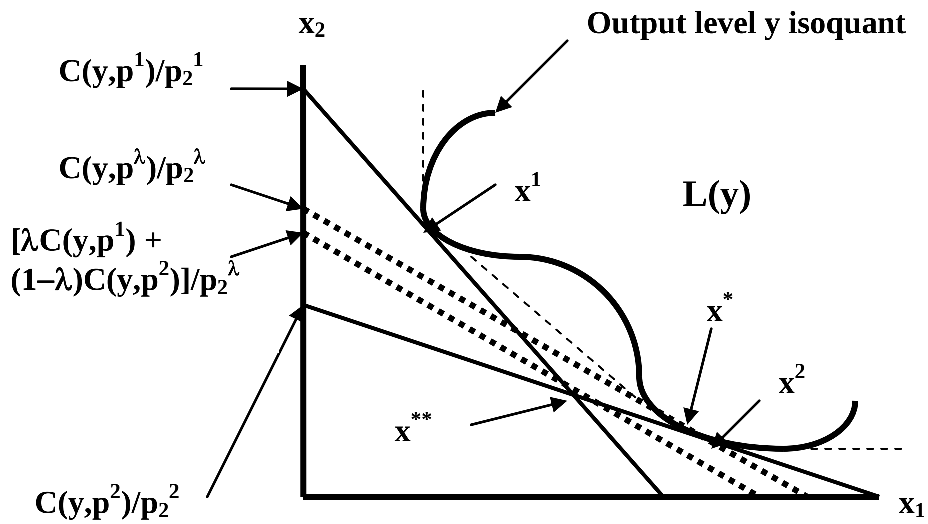
$$\begin{aligned} C(y,\lambda p^1+(1-\lambda)p^2) &\equiv \min_x \{[\lambda p^1+(1-\lambda)p^2]^Tx : f(x) \ge y \,;\, x \ge 0_N\} \\ &= [\lambda p^1 + (1-\lambda)p^2]^Tx^* \qquad \text{where } x^* \ge 0_N \text{ and } f(x^*) \ge y \\ &= \lambda p^{1T}x^* + (1-\lambda)p^{2T}x^* \\ &\ge \lambda \min_x \{p^{1T}x : f(x) \ge y \,;\, x \ge 0_N\} + (1-\lambda)p^{2T}x^* \end{aligned}$$

since x^* is feasible for the cost minimization problem that uses the price vector p^1 and using also $\lambda > 0$

$= \lambda C(y,p^1) + (1-\lambda)p^{2T}x^*$ using the definition of $C(y,p^1)$
$\geq \lambda C(y,p^1) + (1-\lambda)\min_x \{p^{2T}x : f(x) \geq y ; x \geq 0_N\}$
since x^* is feasible for the cost minimization problem that uses the price vector p^2 and using also $1-\lambda > 0$
$= \lambda C(y,p^1) + (1-\lambda)C(y,p^2)$ using the definition of $C(y,p^2)$.

Figure 1 below illustrates why this concavity property holds.

Figure 1: The Concavity in Prices Property of the Cost Function



In Figure 1, the isocost line $\{x: p^{1T}x = C(y,p^1)\}$ is tangent to the production possibilities set $L(y) \equiv \{x: f(x) \geq y, x \geq 0_N\}$ at the point x^1 and the isocost line $\{x: p^{2T}x = C(y,p^2)\}$ is tangent to the production possibilities set $L(y)$ at the point x^2. Note that the point x^{**} belongs to both of these isocost lines. Thus x^{**} will belong to any weighted average of the two isocost lines. The λ and $1-\lambda$ weighted average isocost line is the set $\{x: [\lambda p^1+(1-\lambda)p^2]^T x = \lambda C(y,p^1) + (1-\lambda)C(y,p^2)\}$ and this set is the dotted line through x^{**} in Figure 1. This dotted line lies *below*[10] the parallel dotted line that is just tangent to $L(y)$, which is the isocost line $\{x: [\lambda p^1+(1-\lambda)p^2]^T x = [\lambda p^1+(1-\lambda)p^2]^T x^* = C(y,\lambda p^1+(1-\lambda)p^2)\}$ and it is this fact that gives us the concavity inequality (4).

Proof of Property 5: Since $C(y,p)$ is a concave function of p defined over the open set of p's, $\Omega \equiv \{p: p >> 0_N\}$, it follows that $C(y,p)$ is also continuous in p over this domain of definition set for each fixed $y \in$ Range f.[11]

Proof of Property 6: Let $p >> 0_N$, $y^1 \in$ Range f, $y^2 \in$ Range f, $y^1 < y^2$. Then

[10] It can happen that the two dotted lines coincide.
[11] See Fenchel (1953; 75) or Rockafellar (1970; 82).

$C(y^2,p) \equiv \min_x \{p^Tx : f(x) \geq y^2 ; x \geq 0_N\}$
$\geq \min_x \{p^Tx : f(x) \geq y^1 ; x \geq 0_N\}$
since if $y^1 < y^2$, the set $\{x : f(x) \geq y^2\}$ is a subset of the set $\{x : f(x) \geq y^1\}$ and the minimum of a linear function over a bigger set cannot increase
$\equiv C(y^1,p)$.

Proof of Property 7: The proof is rather technical and may be found in Diewert (1993; 113-114). Q.E.D.

Problems

1. In industrial organization,[12] it once was fairly common to assume that a firm's cost function had the following linear functional form: $C(y,p) \equiv \alpha + \beta^Tp + \gamma y$ where α and γ are scalar parameters and β is a vector of parameters to be estimated econometrically. What are sufficient conditions on these N+2 parameters for this cost function to satisfy properties 1 to 7 above? Is the resulting cost function very realistic?

2. Suppose a producer's production function, f(x), defined for $x \in S$ where $S \equiv \{x: x \geq 0_N\}$ satisfies the following conditions:
(i) f is continuous over S;
(ii) $f(x) > 0$ if $x >> 0_N$ and
(iii) f is positively linearly homogeneous over S; i.e., for every $x \geq 0_N$ and $\lambda > 0$, $f(\lambda x) = \lambda f(x)$.
Define the producer's *unit cost function* c(p) for $p >> 0_N$ as follows:
(iv) $c(p) \equiv C(1,p) \equiv \min_x \{p^Tx : f(x) \geq 1 ; x \geq 0_N\}$;
i.e., c(p) is the minimum cost of producing one unit of output if the producer faces the positive input price vector p. For $y > 0$ and $p >> 0_N$, *show that*
(v) $C(y,p) = c(p)y$.
Note: A production function f that satisfies property (iii) is said to exhibit *constant returns to scale*. The interpretation of (v) is that if a production function exhibits constant returns to scale, then total cost is equal to unit cost times the output level.[13]

3. Shephard (1953; 4) defined a production function F to be *homothetic* if it could be written as
(i) $F(x) = g[f(x)]$; $x \geq 0_N$
where f satisfies conditions (i)-(iii) in Problem 2 above and g(z), defined for all $z \geq 0$, satisfies the following regularity conditions:
(ii) g(z) is positive if $z > 0$;
(iii) g is a continuous function of one variable and
(iv) g is monotonically increasing; i.e., if $0 \leq z^1 < z^2$, then $g(z^1) < g(z^2)$.
Let C(y,p) be the cost function that corresponds to F(x). *Show that* under the above assumptions, for $y > 0$ and $p >> 0_N$, we have
(v) $C(y,p) = g^{-1}(y)c(p)$

[12] For example, see Walters (1961).
[13] We will study the unit cost function in more detail in Section 6 below.

where c(p) is the unit cost function that corresponds to the linearly homogeneous f and g^{-1} is the inverse function for g; i.e., $g^{-1}[g(z)] = z$ for all $z \geq 0$. Note that $g^{-1}(y)$ is a monotonically increasing continuous function of one variable.

3. The Duality Between Cost and Production Functions

The material in the previous section shows how the cost function can be determined from a knowledge of the production function. We now ask whether a knowledge of the cost function is sufficient to determine the underlying production function. The answer to this question is *yes*, but with some qualifications.

To see how we might use a given cost function (satisfying the 7 regularity conditions listed in the previous section) to determine the production function that generated it, pick an arbitrary feasible output level $y > 0$ and an arbitrary vector of positive prices, $p^1 \gg 0_N$ and use the given cost function C to define the following *isocost surface*: $\{x: p^{1T}x = C(y,p^1)\}$. This isocost surface must be tangent to the set of feasible input combinations x that can produce at least output level y, which is the upper level set, $L(y) \equiv \{x: f(x) \geq y;\ x \geq 0_N\}$. It can be seen that this isocost surface and the set lying above it must contain the upper level set L(y); i.e., the following *halfspace* $M(y,p^1)$, contains L(y):

(7) $M(y,p^1) \equiv \{x: p^{1T}x \geq C(y,p^1)\}$.

Pick another positive vector of prices, $p^2 \gg 0_N$ and it can be seen, repeating the above argument, that the halfspace $M(y,p^2) \equiv \{x: p^{2T}x \geq C(y,p^2)\}$ must also contain the upper level set L(y). Thus L(y) must belong to the intersection of the two halfspaces $M(y,p^1)$ and $M(y,p^2)$. Continuing to argue along these lines, it can be seen that L(y) must be contained in the following set, which is the intersection of all of the supporting halfspaces to L(y):

(8) $M(y) \equiv \bigcap_{p \gg 0_N} M(y,p)$.

Note that M(y) is defined using just the given cost function, C(y,p). Note also that since each of the sets in the intersection, M(y,p), is a convex set, then M(y) is also a convex set. Since L(y) is a subset of each M(y,p), it must be the case that L(y) is also a subset of M(y); i.e., we have

(9) $L(y) \subset M(y)$.

Is it the case that L(y) is equal to M(y)? In general, the answer is *no*; M(y) forms an *outer approximation* to the true production possibilities set L(y). To see why this is, see Figure 1 above. The boundary of the set M(y) partly coincides with the boundary of L(y) but it encloses a bigger set: the backward bending parts of the isoquant $\{x: f(x) = y\}$ are replaced by the dashed lines that are parallel to the x_1 axis and the x_2 axis and the inward bending part of the true isoquant is replaced by the dashed line that is tangent to the two regions where the boundary of M(y) coincides with the boundary of L(y). However, if the producer is a price taker in input markets, then it can be seen that *we will never observe* *the producer's nonconvex portions or backwards bending parts of the isoquant*.[14] Thus under the assumption of competitive behavior in input markets, there is no loss of generality in assuming that the producer's production function is *nondecreasing* (this will eliminate the backward bending isoquants) or in assuming that the upper level sets of the production function are convex sets (this will eliminate the nonconvex portions of the upper level sets). A function has convex upper level sets if and only if it is *quasiconcave*.[15]

Putting the above material together, we see that conditions on the production function f(x) that are necessary for the sets M(y) and L(y) to coincide are:

(10) f(x) is defined for $x \geq 0_N$ and is continuous from above[16] over this domain of definition set;
(11) f is nondecreasing and
(12) f is quasiconcave.

Theorem 2: *Shephard Duality Theorem*:[17] If f satisfies (10)-(12), then the cost function C defined by (1) satisfies the properties listed in Theorem 1 above and the upper level sets M(y) defined by (8) using only the cost function coincide with the upper level sets L(y) defined using the production function; i.e., under these regularity conditions, the production function and the cost function determine each other.

We consider how an explicit formula for the production function in terms of the cost function can be obtained. Suppose we have a given cost function, C(y,p), and we are given a strictly positive input vector, $x >> 0_N$, and we ask what is the maximum output that this x can produce. It can be seen that

(13) $f(x) = \max_y \{y: x \in M(y)\}$
$= \max_y \{y: C(y,p) \leq p^T x \text{ for every } p >> 0_N\}$ using definitions (7) and (8).
$= \max_y \{y: C(y,p) \leq 1 \text{ for every } p >> 0_N \text{ such that } p^T x = 1\}$

where the last equality follows using the fact that C(y,p) is linearly homogeneous in p as is the function $p^T x$ and hence we can normalize the prices so that $p^T x = 1$.

---

14 Hotelling (1935; 74) made this point many years ago.

15 f is a *quasiconcave function* defined over a convex subset S of R^N if f has the following property: $x^1 \in S$, $x^2 \in S$, $0 < \lambda < 1$ implies $f(\lambda x^1 + (1-\lambda)x^2) \geq \min \{f(x^1), f(x^2)\}$; see Fenchel (1953; 117).

16 Since each of the sets M(y,p) in the intersection set M(y) defined by (8) are closed, it can be shown that M(y) is also a closed set. Hence if M(y) is to coincide with L(y), we need the upper level sets of f to be closed sets and this will hold if and only if f is continuous from above.

17 Shephard (1953) (1970) was the pioneer in establishing various duality theorems between cost and production functions. See also Samuelson (1953), Uzawa (1964), McFadden (1966) (1978), Diewert (1971) (1974a; 116-118) and Blackorby, Primont and Russell (1978) for various duality theorems under alternative regularity conditions. Our exposition follows that of Diewert (1993; 107-117). These duality theorems are global in nature; i.e., the production and cost functions satisfy their appropriate regularity conditions over their entire domains of definition. However, it is also possible to develop duality theorems that are local rather than global; see Blackorby and Diewert (1979).

We now consider the continuity properties of C(y,p) with respect to p. We have defined C(y,p) for all strictly positive price vectors p and since this domain of definition set is open, we know that C(y,p) is also continuous in p over this set, using the concavity in prices property of C. We would like to extend the domain of definition of C(y,p) from the strictly positive orthant of prices, $\Omega \equiv \{p: p >> 0_N\}$, to the nonnegative orthant, Clo $\Omega \equiv \{p: p \geq 0_N\}$, which is the closure of Ω. It turns out that it is possible to do this if we make use of some theorems in convex analysis.

Theorem 3: *Continuity from above of a concave function using the Fenchel closure operation*: Fenchel (1953; 78): Let f(x) be a concave function of N variables defined over the open convex subset S of R^N. Then there exists a unique extension of f to Clo S, the closure of S, which is concave and continuous from above.

Proof: Using one of Fenchel's (1953; 57) characterizations of concavity, the *hypograph* of f, $H \equiv \{(y,x): y \leq f(x); x \in S\}$, is a convex set in R^{N+1}. Hence the closure of H, Clo H, is also a convex set. Hence the following function f^* defined over Clo S is also a concave function:

$$(14)\quad f^*(x) \equiv \max_y \{y: (y,x) \in \text{Clo } H\}; \qquad x \in \text{Clo } S.$$
$$= f(x) \qquad \text{for } x \in S.$$

Since Clo H is a closed set, it turns out that f^* is continuous from above. Q.E.D.

To see that the extension function f^* need not be continuous, consider the following *example,* where the domain of definition set is $S \equiv \{(x_1,x_2); x_2 \in R^1, x_1 \geq x_2^2\}$ in R^2:

$$(15)\quad f(x_1,x_2) \equiv -x_2^2/x_1 \text{ if } x_2 \neq 0, x_1 \geq x_2^2;$$
$$\equiv 0 \qquad \text{if } x_1 = 0 \text{ and } x_2 = 0.$$

It is possible to show that f is concave and hence continuous over the interior of S; see problem 5 below. However, it can be shown that f is not continuous at (0,0). Let (x_1,x_2) approach (0,0) along the line $x_1 = x_2 > 0$. Then

$$(16)\quad \lim_{x_1 \to 0} f(x_1,x_2) = \lim_{x_1 \to 0} [-x_1^2/x_1] = \lim_{x_1 \to 0} [-x_1] = 0.$$

Now let (x_1,x_2) approach (0,0) along the parabolic path $x_2 > 0$ and $x_1 = x_2^2$. Then

$$(17)\quad \lim_{x_2 \to 0;\ x_1 = x_2^2} f(x_1,x_2) = \lim_{x_2 \to 0} -x_2^2/x_2^2 = -1.$$

Thus f is not continuous at (0,0). It can be verified that restricting f to Int S and then extending f to the closure of S (which is S) leads to the same f^* as is defined by (15). Thus the Fenchel closure operation does not always result in a continuous concave function.

Theorem 4 below states sufficient conditions for the Fenchel closure of a concave function defined over an open domain of definition set to be continuous over the closure of the original domain of definition. Fortunately, the hypotheses of this Theorem are weak enough to cover most economic applications. Before stating the Theorem, we need an additional definition.

Definition: A set S in R^N is a *polyhedral set* iff S is equal to the intersection of a *finite* number of halfspaces.

Theorem 4: *Continuity of a concave function using the Fenchel closure operation*; Gale, Klee and Rockafellar (1968), Rockafellar (1970; 85): Let f be a concave function of N variables defined over an open convex polyhedral set S. Suppose f is bounded from below over every bounded subset of S. Then the Fenchel closure extension of f to the closure of S results in a continuous concave function defined over Clo S.

The proof of this result is too involved to reproduce here but we can now apply this result.

Applying Theorem 4, extend the domain of definition of C(y,p) from strictly positive price vectors p to nonnegative price vectors using the Fenchel closure operation and hence C(y,p) will be continuous and concave in p over the set $\{p: p \geq 0_N\}$ for each y in the interval of feasible outputs.[18]

Now return to the problem where we have a given cost function, C(y,p), we are given a strictly positive input vector, $x \gg 0_N$, and we ask what is the maximum output that this x can produce. Repeating the analysis in (13), we have

(18) $f(x) = \max_y \{y: x \in M(y)\}$
$= \max_y \{y: C(y,p) \leq p^T x \text{ for every } p \gg 0_N\}$ using definitions (7) and (8).
$= \max_y \{y: C(y,p) \leq 1 \text{ for every } p \gg 0_N \text{ such that } p^T x = 1\}$
where we have used the linear homogeneity in prices property of C
$= \max_y \{y: C(y,p) \leq 1 \text{ for every } p \geq 0_N \text{ such that } p^T x = 1\}$
where we have extended the domain of definition of C(y,p) to nonnegative prices from positive prices and used the continuity of the extension function over the set of nonnegative prices
$= \max_y \{y: G(y,x) \leq 1\}$

where the function G(y,x) is defined as follows:

(19) $G(y,x) \equiv \max_p \{C(y,p): p \geq 0_N \text{ and } p^T x = 1\}$.

Note that the maximum in (19) will exist since C(y,p) is continuous in p and the feasible region for the maximization problem, $\{p: p \geq 0_N \text{ and } p^T x = 1\}$, is a closed and bounded

[18] If $f(0_N) = 0$ and f(x) tends to plus infinity as the components of x tend to plus infinity, then the feasible y set will be $y \geq 0$ and C(y,p) will be defined for all $y \geq 0$ and $p \geq 0_N$.

set.[19] Property 7 on the cost function C(y,p) will imply that the maximum in the last line of (18) will exist. Property 6 on the cost function will imply that for fixed x, G(y,x) is nondecreasing in y. Typically, G(y,x) will be continuous in y for a fixed x and so the maximum y that solves (18) will be the y^* that satisfies the following equation:[20]

(20) $G(y^*,x) = 1$.

Thus (19) and (20) implicitly define the production function $y^* = f(x)$ in terms of the cost function C.

Problems

4. Show that the $f(x_1,x_2)$ defined by (15) above is a concave function over the interior of the domain of definition set S. You do not have to show that S is a convex set.

5. In the case where the technology is subject to constant returns to scale, the cost function has the following form: C(y,p) = yc(p) where c(p) is a unit cost function. For $x >> 0_N$, define the function g(x) as follows:
(i) $g(x) \equiv \max_p \{c(p): p^Tx = 1; p \geq 0_N\}$.
Show that in this constant returns to scale case, the function G(y,x) defined by (19) reduces to
(ii) G(y,x) = yg(x).
Show that in this constant returns to scale case, the production function that is dual to the cost function has the following explicit formula for $x >> 0_N$:
(iii) f(x) = 1/g(x).

6. Let $x \geq 0$ be input (a scalar number) and let $y = f(x) \geq 0$ be the maximum output that could be produced by input x, where f is the production function. Suppose that f is defined as the following *step function*:

(i) $f(x) \equiv 0$ for $0 \leq x < 1$;
$\equiv 1$ for $1 \leq x < 2$;
$\equiv 2$ for $2 \leq x < 3$;

and so on. Thus the technology cannot produce fractional units of output and it takes one full unit of input to produce each unit of output. It can be verified that this production function is continuous from above.
(a) Calculate the cost function C(y,1) that corresponds to this production function; i.e., set the input price equal to one and try to determine the corresponding total cost function C(y,1). It will turn out that this cost function is continuous from below in y.
(b) Graph both the production function y = f(x) and the cost function $c(y) \equiv C(y,1)$.

---

[19] Here is where we use the assumption that $x >> 0_N$ in order to obtain the boundedness of this set.
[20] This method for constructing the production function from the cost function may be found in Diewert (1974a; 119).

7. Suppose that a producer's cost function is defined as follows for $y \geq 0$, $p_1 > 0$ and $p_2 > 0$:

(i) $C(y,p_1,p_2) \equiv [b_{11}p_1 + 2b_{12}(p_1p_2)^{1/2} + b_{22}p_2]y$

where the b_{ij} parameters are all positive.
(a) Show that this cost function is concave in the input prices p_1,p_2. *Note*: this is the two input case of the Generalized Leontief cost function defined by Diewert (1971).
(b) Calculate an explicit functional form for the corresponding production function $f(x_1,x_2)$ where we assume that $x_1 > 0$ and $x_2 > 0$. *Hint*: This part of the problem is not completely straightforward. You will obtain a quadratic equation but which root is the right one?

4. The Derivative Property of the Cost Function

Theorem 2, the Shephard Duality Theorem, is of mainly academic interest: if the production function f satisfies properties (10)-(12), then the corresponding cost function C defined by (1) satisfies the properties listed in Theorem 1 above and moreover completely determines the production function. However, it is the next property of the cost function that makes duality theory so useful in applied economics.

Theorem 5: Shephard's (1953; 11) *Lemma*: If the cost function C(y,p) satisfies the properties listed in Theorem 1 above and in addition is once differentiable with respect to the components of input prices at the point (y^*,p^*) where y^* is in the range of the production function f and $p^* >> 0_N$, then

(21) $x^* = \nabla_p C(y^*,p^*)$

where $\nabla_p C(y^*,p^*)$ is the vector of first order partial derivatives of cost with respect to input prices, $[\partial C(y^*,p^*)/\partial p_1,...,\partial C(y^*,p^*)/\partial p_N]^T$, and x^* is any solution to the cost minimization problem

(22) $\min_x \{p^{*T}x: f(x) \geq y^*\} \equiv C(y^*,p^*)$.

Under these differentiability hypotheses, it turns out that the x^* solution to (22) is unique.

Proof: Let x^* be any solution to the cost minimization problem (22). Since x^* is feasible for the cost minimization problem when the input price vector is changed to an arbitrary $p >> 0_N$, it follows that

(23) $p^Tx^* \geq C(y^*,p)$ for every $p >> 0_N$.

Since x^* is a solution to the cost minimization problem (22) when $p = p^*$, we must have

(24) $p^{*T}x^* = C(y^*,p^*)$.

But (23) and (24) imply that the function of N variables, $g(p) \equiv p^T x^* - C(y^*,p)$ is nonnegative for all $p >> 0_N$ with $g(p^*) = 0$. Hence, g(p) attains a global minimum at $p = p^*$ and since g(p) is differentiable with respect to the input prices p at this point, the following first order necessary conditions for a minimum must hold at this point:

$$(25)\ \nabla_p\, g(p^*) = x^* - \nabla_p C(y^*,p^*) = 0_N.$$

Now note that (25) is equivalent to (21). If x^{**} is any other solution to the cost minimization problem (22), then repeat the above argument to show that

$$\begin{aligned}(26)\ x^{**} &= \nabla_p C(y^*,p^*)\\ &= x^*\end{aligned}$$

where the second equality follows using (25). Hence $x^{**} = x^*$ and the solution to (22) is unique. Q.E.D.

The above result has the following implication: postulate a differentiable functional form for the cost function C(y,p) that satisfies the regularity conditions listed in Theorem 1 above. Then differentiating C(y,p) with respect to the components of the input price vector p generates the firm's system of cost minimizing input demand functions, $x(y,p) \equiv \nabla_p C(y,p)$.

Shephard (1953) was the first person to establish the above result starting with just a cost function satisfying the appropriate regularity conditions.[21] However, Hotelling (1932; 594) stated a version of the result in the context of profit functions and Hicks (1946; 331) and Samuelson (1953; 15-16) established the result starting with a differentiable utility or production function.

One application of the above result is its use as an aid in generating systems of cost minimizing input demand functions that are linear in the parameters that characterize the technology. For example, suppose that the cost function had the following *Generalized Leontief functional form*:[22]

$$(27)\ C(y,p) \equiv \sum_{i=1}^N \sum_{k=1}^N b_{ik}\, p_i^{1/2}\, p_k^{1/2}\, y\ ; \qquad b_{ik} = b_{ki} \text{ for } 1 \le i < j \le N$$

where the N(N+1)/2 independent b_{ik} parameters are all nonnegative. With these nonnegativity restrictions, it can be verified that the C(y,p) defined by (27) satisfies properties 1 to 7 listed in Theorem 1.[23] Applying Shephard's Lemma shows that the system of cost minimizing input demand functions that correspond to this functional form are given by:

[21] This is why Diewert (1974a; 112) called the result Shephard's Lemma. See also Fenchel (1953; 104). We have used the technique of proof used by McKenzie (1956-57).
[22] See Diewert (1971).
[23] Using problem 7 above, it can be seen that if the b_{ik} are nonnegative and y is positive, then the functions $b_{ik}\, p_i^{1/2} p_k^{1/2}\, y$ are concave in the components of p. Hence, since a sum of concave functions is concave, it can be seen that the C(y,p) defined by (27) is concave in the components of p.

(28) $x_i(y,p) = \partial C(y,p)/\partial p_i = \sum_{k=1}^{N} b_{ik} (p_k/p_i)^{1/2} y$; $\quad i = 1,2,\ldots,N.$

Errors can be added to the system of equations (28) and the parameters b_{ik} can be estimated using linear regression techniques if we have time series or cross sectional data on output, inputs and input prices.[24] If all of the b_{ij} equal zero for $i \neq j$, then the demand functions become:

(29) $x_i(y,p) = \partial C(y,p)/\partial p_i = b_{ii}\, y$; $\quad i = 1,2,\ldots,N.$

Note that input prices do not appear in the system of input demand functions defined by (29) so that input quantities do not respond to changes in the relative prices of inputs. The corresponding production function is known as the Leontief (1941) production function.[25] Hence, it can be seen that the production function that corresponds to (28) is a generalization of this production function.

We will consider additional functional forms for a cost function in subsequent sections.

5. The Comparative Statics Properties of Input Demand Functions

Before we develop the main result in this section, it will be useful to establish some results about the derivatives of a twice continuously differentiable linearly homogeneous function of N variables. We say that f(x), defined for $x >> 0_N$ is *positively homogeneous of degree* α iff f has the following property:

(30) $f(\lambda x) = \lambda^{\alpha} f(x)$ $\quad$ for all $x >> 0_N$ and $\lambda > 0$.

A special case of the above definition occurs when the number α in the above definition equals 1. In this case, we say that f is (positively) *linearly homogeneous*[26] iff

(31) $f(\lambda x) = \lambda f(x)$ $\quad$ for all $x >> 0_N$ and $\lambda > 0$.

Theorem 6: *Euler's Theorems on Differentiable Homogeneous Functions*: Let f(x) be a (positively) linearly homogeneous function of N variables, defined for $x >> 0_N$. Part 1: If the first order partial derivatives of f exist, then the first order partial derivatives of f satisfy the following equation:

---

[24] Note that b_{12} will appear in the first input demand equation and in the second as well using the cross equation symmetry condition, $b_{21} = b_{12}$. There are $N(N-1)/2$ such cross equation symmetry conditions and we could test for their validity or impose them in order to save degrees of freedom. The nonnegativity restrictions that ensure global concavity of C(y,p) in p can be imposed if we replace each parameter b_{ik} by a squared parameter, $(a_{ik})^2$. However, the resulting system of estimating equations is no longer linear in the unknown parameters.

[25] The Leontief production function can be defined as $f(x_1,\ldots,x_N) \equiv \min_i \{x_i/b_{ii} : i = 1,\ldots,N\}$. It is also known as the no substitution production function. Note that this production function is not differentiable even though its cost function is differentiable.

[26] Usually in economics, we omit the adjective "positively" but it is understood that the λ which appears in definitions (30) and (31) is restricted to be positive.

(32) $f(x) = \sum_{n=1}^{N} x_n \, \partial f(x_1,\ldots,x_N)/\partial x_n = x^T\nabla f(x)$ for all $x >> 0_N$.

Part 2: If the second order partial derivatives of f exist, then they satisfy the following equations:

(33) $\sum_{k=1}^{N} [\partial^2 f(x_1,\ldots,x_N)/\partial x_n \partial x_k]x_k = 0$ for all $x >> 0_N$ and $n = 1,\ldots,N$.

The N equations in (33) can be written using matrix notation in a much more compact form as follows:

(34) $\nabla^2 f(x)x = 0_N$ for all $x >> 0_N$.

Proof of Part 1: Let $x >> 0_N$ and $\lambda > 0$. Differentiating both sides of (31) with respect to λ leads to the following equation using the composite function chain rule:

$$(35)\ f(x) = \sum_{n=1}^{N} [\partial f(\lambda x_1,\ldots,\lambda x_N)/\partial(\lambda x_n)][\partial(\lambda x_n)/\partial\lambda]$$
$$= \sum_{n=1}^{N} [\partial f(\lambda x_1,\ldots,\lambda x_N)/\partial(\lambda x_n)]x_n .$$

Now evaluate (35) at $\lambda = 1$ and we obtain (32).

Proof of Part 2: Let $x >> 0_N$ and $\lambda > 0$. For $n = 1,\ldots,N$, differentiate both sides of (31) with respect to x_n and we obtain the following N equations:

(36) $f_n(\lambda x_1,\ldots,\lambda x_N)\partial(\lambda x_n)/\partial x_n = \lambda f_n(x_1,\ldots,x_N)$ for $n = 1,\ldots,N$ or
$f_n(\lambda x_1,\ldots,\lambda x_N)\lambda = \lambda f_n(x_1,\ldots,x_N)$ for $n = 1,\ldots,N$ or
$f_n(\lambda x_1,\ldots,\lambda x_N) = f_n(x_1,\ldots,x_N)$ for $n = 1,\ldots,N$

where the nth first order partial derivative function is defined as $f_n(x_1,\ldots,x_N) \equiv \partial f(x_1,\ldots,x_N)/\partial x_n$ for $n = 1,\ldots,N$.[27] Now differentiate both sides of the last set of equations in (36) with respect to λ and we obtain the following N equations:

$$(37)\ 0 = \sum_{k=1}^{N} [\partial f_n(\lambda x_1,\ldots,\lambda x_N)/\partial x_k][\partial(\lambda x_k)/\partial\lambda] \quad \text{for } n = 1,\ldots,N$$
$$= \sum_{k=1}^{N} [\partial f_n(\lambda x_1,\ldots,\lambda x_N)/\partial x_k]x_k .$$

Now evaluate (37) at $\lambda = 1$ and we obtain the N equations (33). Q.E.D.

The above results can be applied to the cost function, C(y,p). From Theorem 1, C(y,p) is linearly homogeneous in p. Hence by part 2 of Euler's Theorem, if the second order partial derivatives of the cost function with respect to the components of the input price vector p exist, then these derivatives satisfy the following restrictions:

[27] Using definition (30) for the case where $\alpha = 0$, it can be seen that the last set of equations in (36) shows that the first order partial derivative functions of a linearly homogenous function are homogeneous of degree 0.

(38) $\nabla^2_{pp}C(y,p)p = 0_N$.

Theorem 7: Diewert (1993; 148-150): Suppose the cost function C(y,p) satisfies the properties listed in Theorem 1 and in addition is twice continuously differentiable with respect to the components of its input price vector at some point, (y,p). Then the system of cost minimizing input demand equations, $x(y,p) \equiv [x_1(y,p),\ldots,x_N(y,p)]^T$, exists at this point and these input demand functions are once continuously differentiable. Form the N by N matrix of input demand derivatives with respect to input prices, $B \equiv [\partial x_i(y,p)/\partial p_j]$, which has ij element equal to $\partial x_i(y,p)/\partial p_j$. Then the matrix B has the following properties:

(39) $B = B^T$ so that $\partial x_i(y,p)/\partial p_k = \partial x_k(y,p)/\partial p_i$ for all $i \neq k$;[28]
(40) B is negative semidefinite[29] and
(41) $Bp = 0_N$.[30]

Proof: Shephard's Lemma implies that the firm's system of cost minimizing input demand equations, $x(y,p) \equiv [x_1(y,p),\ldots,x_N(y,p)]^T$, exists and is equal to

(42) $x(y,p) = \nabla_p C(y,p)$.

Differentiating both sides of (42) with respect to the components of p gives us

(43) $B \equiv [\partial x_i(y,p)/\partial p_k] = \nabla^2_{pp}C(y,p)$.

Now property (39) follows from Young's Theorem in calculus. Property (40) follows from (43) and the fact that C(y,p) is concave in p and the fourth characterization of concavity. Finally, property (41) follows from the fact that the cost function is linearly homogeneous in p and hence (38) holds. Q.E.D.

Note that property (40) implies the following properties on the input demand functions:

(44) $\partial x_n(y,p)/\partial p_n \leq 0$ for $n = 1,\ldots,N$.

Property (44) means that input demand curves cannot be upward sloping.

If the cost function is also differentiable with respect to the output variable y, then we can deduce an additional property about the first order derivatives of the input demand

[28] These are the Hicks (1946; 311) and Samuelson (1947; 69) *symmetry restrictions*. Hotelling (1932; 549) obtained analogues to these symmetry conditions in the profit function context.
[29] Hicks (1946; 311) and Samuelson (1947; 69) also obtained versions of this result by starting with the production (or utility) function f(x), assuming that the first order conditions for solving the cost minimization problem held and that the strong second order sufficient conditions for the primal cost minimization problem also held. Thus using duality theory, we obtain the same results under weaker regularity conditions.
[30] Hicks (1946; 331) and Samuelson (1947; 69) also obtained this result using their primal technique.

functions. The linear homogeneity property of C(y,p) in p implies that the following equation holds for all $\lambda > 0$:

$$(45)\ C(y,\lambda p) = \lambda C(y,p) \qquad \text{for all } \lambda > 0 \text{ and } p >> 0_N.$$

Partially differentiating both sides of (45) with respect to y leads to the following equation:

$$(46)\ \partial C(y,\lambda p)/\partial y = \lambda \partial C(y,p)/\partial y \qquad \text{for all } \lambda > 0 \text{ and } p >> 0_N.$$

But (46) implies that the function $\partial C(y,p)/\partial y$ is linearly homogeneous in p and hence part 1 of Euler's Theorem applied to this function gives us the following equation:

$$(47)\ \partial C(y,p)/\partial y = \sum_{n=1}^{N} p_n \partial^2 C(y,p)/\partial y \partial p_n = p^T \nabla^2_{yp} C(y,p).$$

But using (42), it can be seen that (47) is equivalent to the following equation:[31]

$$(48)\ \partial C(y,p)/\partial y = \sum_{n=1}^{N} p_n \partial x_n(y,p)/\partial y.$$

Problems

8. For $i \neq k$, the inputs i and k are said to be substitutes if $\partial x_i(y,p)/\partial p_k = \partial x_k(y,p)/\partial p_i > 0$, unrelated if $\partial x_i(y,p)/\partial p_k = \partial x_k(y,p)/\partial p_i = 0$ [32] and complements if $\partial x_i(y,p)/\partial p_k = \partial x_k(y,p)/\partial p_i < 0$. (a) If N = 2, show that the two inputs cannot be complements. (b) If N = 2 and $\partial x_1(y,p)/\partial p_1 = 0$, then show that all of the remaining input demand price derivatives are equal to 0; i.e., show that $\partial x_1(y,p)/\partial p_2 = \partial x_2(y,p)/\partial p_1 = \partial x_2(y,p)/\partial p_2 = 0$. (c) If N = 3, show that at most one pair of inputs can be complements.[33]

9. Let $N \geq 3$ and suppose that $\partial x_1(y,p)/\partial p_1 = 0$. Then show that $\partial x_1(y,p)/\partial p_n = 0$ as well for n = 2,3,…,N. *Hint*: You may need to use the definition of negative semidefiniteness in a strategic way. This problem shows that if the own input elasticity of demand for an input is 0, then that input is unrelated to all other inputs.

10. Recall the definition (27) of the Generalized Leontief cost function where the parameters b_{ij} were all assumed to be nonnegative. Show that under these nonnegativity restrictions, every input pair is either unrelated or substitutes. *Hint*: Simply calculate $\partial^2 C(y,p)/\partial p_i \partial p_k$ for $i \neq k$ and look at the resulting formula. *Comment*: This result shows that if we impose the nonnegativity conditions $b_{ik} \geq 0$ for $i \neq j$ on this functional form in order to ensure that it is globally concave in prices, then we have a priori ruled out any form of complementarity between the inputs. This means if the number of inputs N is

[31] This method of deriving these restrictions is due to Diewert (1993; 150) but these restrictions were originally derived by Samuelson (1947; 66) using his primal cost minimization method.
[32] Pollak (1969; 67) used the term "unrelated" in a similar context.
[33] This result is due to Hicks (1946; 311-312): "It follows at once from Rule (5) that, while it is possible for all other goods consumed to be substitutes for x_1, it is not possible for them all to be complementary with it."

greater than 2, this nonnegativity restricted functional form cannot be a *flexible functional form*[34] for a cost function; i.e., it cannot attain an arbitrary pattern of demand derivatives that are consistent with microeconomic theory, since the nonnegativity restrictions rule out any form of complementarity.

11. Suppose that a producer's three input production function has the following Cobb Douglas (1928) functional form:

(a) $f(x_1,x_2,x_3) \equiv x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$ where $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 > 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

Let the positive input prices $p_1 > 0$, $p_2 > 0$, $p_3 > 0$ and the positive output level $y > 0$ be given. (i) Calculate the producer's cost function, $C(y,p_1,p_2,p_3)$ along with the three input demand functions, $x_1(y,p_1,p_2,p_3)$, $x_2(y,p_1,p_2,p_3)$ and $x_3(y,p_1,p_2,p_3)$. *Hint*: Use the usual Lagrangian technique for solving constrained minimization problems. You do not need to check the second order conditions for the problem. The positive constant $k \equiv \alpha_1^{-\alpha_1}\alpha_2^{-\alpha_2}\alpha_3^{-\alpha_3}$ will appear in the cost function.

(ii) Calculate the input one demand elasticity with respect to output $[\partial x_1(y,p_1,p_2,p_3)/\partial y][y/x_1(y,p_1,p_2,p_3)]$ and the three input one demand elasticities with respect to input prices $[\partial x_1(y,p_1,p_2,p_3)/\partial p_n][p_n/x_1(y,p_1,p_2,p_3)]$ for n = 1,2,3.

(iii) Show that $-1 < [\partial x_1(y,p_1,p_2,p_3)/\partial p_1][p_1/x_1(y,p_1,p_2,p_3)] < 0$.
(iv) Show that $0 < [\partial x_1(y,p_1,p_2,p_3)/\partial p_2][p_2/x_1(y,p_1,p_2,p_3)] < 1$.
(v) Show that $0 < [\partial x_1(y,p_1,p_2,p_3)/\partial p_3][p_3/x_1(y,p_1,p_2,p_3)] < 1$.

(vi) Can any pair of inputs be complementary if the technology is a three input Cobb Douglas?

Comment: The Cobb Douglas functional form is widely used in macroeconomics and in applied general equilibrium models. However, this problem shows that it is not satisfactory if $N \geq 3$. Even in the $N = 2$ case where analogues to (iii) and (iv) above hold, it can be seen that this functional form is not consistent with technologies where the degree of substitution between inputs is very high or very low.

12. Suppose that the second order partial derivatives with respect to input prices of the cost function $C(y,p)$ exist so that the nth cost minimizing input demand function $x_n(y,p) = \partial C(y,p)/\partial p_n > 0$ exists for $n = 1,\ldots,N$. Define the input n *elasticity of demand* with respect to input price k as follows:

[34] Diewert (1974a; 115) introduced the term "flexible functional form" to describe a functional form for a cost function (or production function) that could approximate an arbitrary cost function (consistent with microeconomic theory) to the second order around any given point. The Generalized Leontief cost function defined by (27) above is flexible for the class of cost functions that are dual to linearly homogeneous production functions if we do not impose any restrictions on the parameters b_{ij}; see Diewert (1971) or Section 8 below for a proof of this fact.

(a) $e_{nk}(y,p) \equiv [\partial x_n(y,p)/\partial p_k][p_k/x_n(y,p)]$ for $n = 1,..,N$ and $k = 1,\ldots,N$.

Show that for each n, $\sum_{k=1}^{N} e_{nk}(y,p) = 0$.

13. Let the producer's cost function be C(y,p), which satisfies the regularity conditions in Theorem 1 and, in addition, is once differentiable with respect to the components of the input price vector p. Then the nth input demand function is $x_n(y,p) \equiv \partial C(y,p)/\partial p_n$ for $n = 1,\ldots,N$. Input n is defined to be *normal* at the point (y,p) if $\partial x_n(y,p)/\partial y = \partial^2 C(y,p)/\partial p_n \partial y > 0$; i.e., if the cost minimizing demand for input n increases as the target output level y increases. On the other hand, input n is defined to be *inferior* at the point (y,p) if $\partial x_n(y,p)/\partial y = \partial^2 C(y,p)/\partial p_n \partial y < 0$. Prove that not all N inputs can be inferior at the point (y,p). *Hint*: Make use of (48).

14. If the production function f dual to the differentiable cost function C(y,p) exhibits *constant returns to scale* so that $f(\lambda x) = \lambda f(x)$ for all $x \geq 0_N$ and all $\lambda > 0$, then show that for each n, the input n elasticity of demand with respect to the output level y is 1; i.e., show that for $n = 1,\ldots,N$, $[\partial x_n(y,p)/\partial y][y/x_n(y,p)] = 1$.

15. Let C(y,p) be a twice continuously differentiable cost function that satisfies the regularity conditions listed in Theorem 1 in section 2 above. By Shephard's Lemma, the input demand functions are given by

(i) $x_n(y,p) = \partial C(y,p)/\partial p_n > 0$; $n = 1,\ldots,N$.

The Allen (1938; 504) Uzawa (1962) *elasticity of substitution* σ_{nk} between inputs n and k is defined as follows:

(ii) $\sigma_{nk}(y,p) \equiv \{C(y,p)\partial^2 C(y,p)/\partial p_n \partial p_k\}/\{[\partial C(y,p)/\partial p_n][\partial C(y,p)/\partial p_k]\}$ $1 \leq n,k \leq N$
$= \{C(y,p)\partial^2 C(y,p)/\partial p_n \partial p_k\}/x_n(y,p)\, x_k(y,p)$ using (i).

Define $\Sigma \equiv [\sigma_{nk}(y,p)]$ as the N by N matrix of elasticities of substitution.
(a) Show that Σ has the following properties:

(iii) $\Sigma = \Sigma^T$;
(iv) Σ is negative semidefinite and
(v) $\Sigma s = 0_N$

where $s \equiv [s_1,\ldots,s_N]^T$ is the vector of cost shares; i.e., $s_n \equiv p_n x_n(y,p)/C(y,p)$ for $n = 1,\ldots,N$. Now define the N by N matrix of cross price elasticities of demand E in a manner analogous to definition (ii) above:

(vi) $E \equiv [e^{nk}]$ $n = 1,\ldots,N$; $k = 1,\ldots,N$
$\equiv [(p_k/x_n)\partial x_n(y,p)/\partial p_k]$
$= [(p_k/x_n)\partial^2 C(y,p)/\partial p_n \partial p_k]$ using (i)
$= \hat{x}^{-1}\nabla^2_{pp}C(y,p)\,\hat{p}$.

(b) Show that $E = \Sigma\hat{s}$ where $\hat{s}$ is an N by N diagonal matrix with the elements of the share vector s running down the main diagonal.

6. The Duality Between Constant Returns to Scale Production Functions and their Unit Cost Functions

In this section, we will add more structure to the production function: we will assume that $f(x)$ is subject to *constant returns to scale* so that $f(\lambda x) = \lambda f(x)$ for every nonnegative input vector $x \geq 0_N$ and nonnegative scalar $\lambda \geq 0$.

In many areas of applied economics, constant returns to scale in production is assumed. Samuelson (1967) justified this assumption as an approximation to reality by using a plant replication argument. He assumed that there was a plant size that minimized average cost and showed if this optimal plant size output level was small relative to the size of the market, then by replicating optimal size plants, the industry production function would approximate a constant returns to scale production function.[35] Thus in this section, we will assume constant returns to scale in production and see what additional properties the resulting cost function must satisfy.

Before we develop a formal duality theorem, it is necessary to prove a useful mathematical result.

Theorem 8: Berge (1963; 208): If f is a positive, linearly homogeneous and quasiconcave function defined over the positive orthant in R^N, Ω, then f is also concave over Ω.

Proof: Let $x^1 \gg 0_N$, $x^2 \gg 0_N$ and $0 < \lambda < 1$. We need to show that:

(49) $f(\lambda x^1 + (1-\lambda)x^2) \geq \lambda f(x^1) + (1-\lambda)f(x^2)$.

Without loss of generality, we can assume $0 < f(x^1) \leq f(x^2)$. Let $\mu > 0$ be the scalar that causes $f(\mu x^2)$ to equal $f(x^1)$. Using the constant returns to scale property of f, μ can be defined as follows:

(50) $\mu \equiv f(x^1)/f(x^2) > 0$.

Points on the line segment joining the point x^1 to μx^2 can be represented by $\alpha x^1 + (1-\alpha)\mu x^2$ where $0 \leq \alpha \leq 1$. The quasiconcavity property of f implies that the following equality holds for all α such that $0 \leq \alpha \leq 1$:

(51) $f(x^1) \leq f(\alpha x^1 + (1-\alpha)\mu x^2)$.

Define $\beta > 0$ as the proportionality factor that deflates the point $\lambda x^1 + (1-\lambda)x^2$ onto the line segment joining the point x^1 to μx^2. Thus we have:

[35] Diewert (1981) elaborated on Samuelson's results.

(52) $\beta[\lambda x^1 + (1-\lambda)x^2] = \alpha x^1 + (1-\alpha)\mu x^2$.

Thus the unknown α and β must be the solution to the following two equations:

(53) $\beta\lambda = \alpha$; $\beta(1-\lambda) = (1-\alpha)\mu$.

The solution to (52) is $\beta = \mu/(1-\lambda+\lambda\mu)$ and $\alpha = \lambda\mu/(1-\lambda+\lambda\mu)$. It is straightforward to show that the solution satisfies $\beta > 0$ and $0 \leq \alpha \leq 1$. Now substitute (52) into (51) and we obtain the following inequality:

(54) $f(x^1) \leq f(\alpha x^1 + (1-\alpha)\mu x^2)$
$= f(\beta[\lambda x^1 + (1-\lambda)x^2])$ using (52)
$= \beta f(\lambda x^1 + (1-\lambda)x^2)$ using the linear homogeneity of f
$= [\mu/(1-\lambda+\lambda\mu)]f(\lambda x^1 + (1-\lambda)x^2)$.

Thus (54) implies:

(55) $f(\lambda x^1 + (1-\lambda)x^2) \geq \mu^{-1}(1-\lambda+\lambda\mu)f(x^1)$
$= \mu^{-1}(1-\lambda)f(x^1) + \lambda f(x^1)$
$= \lambda f(x^1) + (1-\lambda)f(x^2)$ using definition (50).

Q.E.D.

The above result will prove to be useful in what follows. Recall that in Section 2 above, we initially assumed that the production function f(x) only satisfied *continuity from above*. We continue to make this very weak regularity assumption but we now assume that in addition, f satisfies the following *linear homogeneity property*:

(56) $f(\lambda x) = \lambda f(x)$ for all $\lambda \geq 0$ and $x \geq 0_N$.

We also assume that there exists an $x^* > 0_N$ such that $y^* \equiv f(x^*) > 0$; i.e., there exists a nonnegative, nonzero input vector x^* which can produce a positive output. This assumption along with the constant returns to scale assumption (56) means that the technology can produce any positive output level.

Let $y > 0$ and $p >> 0_N$. We can define the *total cost function* that corresponds to our homogeneous production function using definition (1) again; i.e., define C(y,p) as follows:

(57) $C(y,p) \equiv \min_x \{p^Tx : f(x) \geq y ; x \geq 0_N\}$
$= \min_x \{p^Tx : y^{-1}f(x) \geq 1 ; x \geq 0_N\}$
$= \min_x \{p^Tx : f(x/y) \geq 1 ; x \geq 0_N\}$ using (56)
$= \min_x \{yp^T(x/y) : f(x/y) \geq 1 ; x/y \geq 0_N\}$
$= y \min_z \{p^Tz : f(z) \geq 1 ; z \geq 0_N\}$ letting z = x/y
$= y\, c(p)$

where c(p) is the *unit cost function* that corresponds to f, defined as follows:

(58) $c(p) \equiv \min_z \{p^T z : f(z) \geq 1 ; z \geq 0_N\}$.

We can use the input price properties of the total cost function C(y,p) that were implied by Theorem 1 in Section 2 in order to derive the properties of the unit cost function, c(p). Thus Theorem 1 tells us that c(p) is well defined as a minimum for $p >> 0_N$ and it is nonnegative, positively linearly homogeneous, nondecreasing and concave in p over the positive orthant. In fact, the continuity from above property of f along with the assumption that f is linearly homogeneous will imply that $f(0_N) = 0$ and this in turn will imply that $c(p) > 0$ for $p >> 0_N$. Since c(p) is concave over the positive orthant, we can also deduce that it is continuous over this domain of definition. The domain of definition of c(p) can be extended to the nonnegative orthant using the Fenchel closure operation as was done in Section 2. The resulting c(p) will be continuous over the nonnegative orthant. Thus there is no problem in going from the production function to its unit cost function.

Can we use the unit cost function to recover the underlying production function? We can get an *outer approximation* to the true technology using the algebra in Section 2. Let $x > 0_N$ be an arbitrary nonzero, nonnegative input vector. The maximum output y that is consistent with using the outer approximation technology and the input vector x must satisfy the inequalities $yc(p) \leq p^T x$ for every $p > 0_N$. Thus we want the maximum y such that $y \leq p^T x/c(p)$ for every $p > 0_N$. Now the functions $p^T x$ and c(p) are both linearly homogeneous so we can normalize one of these functions and minimize or maximize the remaining function to obtain $y = f^*(x)$, where $f^*(x)$ is the production function that corresponds to the outer approximation technology. If we set $p^T x = 1$, then we want to minimize 1/c(p) subject to the constraint $p^T x = 1$ and so in this case, $f^*(x)$ is defined as follows:

$$
\begin{aligned}
(59)\ f^*(x) &\equiv \min_p \{1/c(p) : p^T x = 1; p \geq 0_N\} \\
&= 1/\max_p \{c(p) : p^T x = 1; p \geq 0_N\}.
\end{aligned}
$$

Note that the maximization problem in (59) is a concave programming problem. On the other hand, we could set c(p) = 1. In this case, $f^*(x)$ is (equivalently) defined as follows:

$$
\begin{aligned}
(60)\ f^*(x) &\equiv \min_p \{p^T x : c(p) = 1; p \geq 0_N\} \\
&= \min_p \{p^T x : c(p) \geq 1; p \geq 0_N\}.
\end{aligned}
$$

In order to recover the original production function, f(x) by using the formulae on the right hand sides of (59) or (60), we need to assume that f is nondecreasing and quasiconcave, as in Section 2. However, using Berge's Theorem 8 above, it can be seen that when f is linearly homogeneous and quasiconcave (and positive) over the positive orthant, then f is also a concave function over the positive orthant. If in addition, f is continuous over the nonnegative orthant, then f will also be concave over the nonnegative orthant. Thus f and c satisfy exactly the same regularity conditions, with respect to x and p respectively if we assume that f is nondecreasing and quasiconcave. Moreover, the underlying technology can be represented by using either the linearly homogeneous production function or its dual unit cost function. Samuelson (1953; 15) and Shephard (1953) were the first to obtain versions of this duality theorem for the homogeneous case.[36]

In the following sections, we will exhibit various explicit functional forms for a linearly homogeneous f or its dual unit cost function.

7. The Constant Elasticity of Substitution Production Function

The Constant Elasticity of Substitution (CES) production function, f(x), is defined as follows:

$$(61)\quad f(x_1,...,x_N) \equiv [\Sigma_{n=1}^N \beta_n x_n^s]^{1/s}$$

where the parameters β_n are positive and s is a parameter which satisfies $s \neq 0$ and the inequality $s \leq 1$. The two input case of this functional form was introduced into the economics literature by Arrow, Chenery, Minhas and Solow (1961; 230).[37] The problems below show that the CES production function is a well behaved constant returns to scale production function which satisfies the regularity conditions that were developed in the previous section, provided that $s \leq 1$.

Problems

16. Let $s \neq 0$ and rewrite the f(x) defined by (61) as $f(x) = \gamma f_s(x)$ where $f_s(x) \equiv [\Sigma_{n=1}^N \beta_n^* x_n^s]^{1/s}$, $\beta_n^* \equiv \beta_n/\Sigma_{i=1}^N \beta_i$ for n = 1,...,N and $\gamma \equiv [\Sigma_{i=1}^N \beta_i]^{1/s}$. Show that $\lim_{s\to 0} \ln f_s(x_1,...,x_N) = \Sigma_{n=1}^N \beta_n^* \ln x_n$. Thus the CES production function defined by (61) tends to a Cobb-Douglas production function as the parameter s tends to 0. *Hint*: Write $\ln f_s(x_1,...,x_N)$ as g(s)/h(s) where $g(s) \equiv \ln[\Sigma_{n=1}^N \beta_n^* x_n^s]$ and $h(s) \equiv s$. Let s tend to 0 and apply l'Hospital's Rule. Note that g(0) = h(0) = 0.

17. Let $\beta^T \equiv [\beta_1,...,\beta_N]$ where $\beta_n > 0$ for n = 1,...,N. Define $\hat{\beta}$ as the N by N diagonal matrix with the elements of the vector β running down the main diagonal. Show that the N by N matrix $-\hat{\beta} + \beta\beta^T$ is a negative semidefinite matrix. *Hint*: Show that the inequality $z^T[-\hat{\beta} + \beta\beta^T]z \leq 0$ for all vectors z is equivalent to the Cauchy-Schwarz inequality $(x^Ty)^2 \leq (x^Tx)(y^Ty)$ with $x \equiv \hat{\beta}^{1/2} 1_N$; $y \equiv \hat{\beta}^{1/2} z$ where 1_N is a vector of ones of dimension N and $\hat{\beta}^{1/2}$ is a diagonal matrix with the positive square roots of the elements of β running down the main diagonal.

---

[36] See also Diewert (1974a; 110-112) for a duality theorem along the present lines.

[37] These authors wrote the CES functional form defined by (61) as $f(x) = \gamma[\Sigma_{n=1}^N \beta_n^* x_n^s]^{1/s}$ where the β_n^* now sum up to one and $\gamma \equiv [\Sigma_{n=1}^N \beta_n]^{1/s}$ is a positive efficiency parameter. They noted that the function of x that is defined by $[\Sigma_{n=1}^N \beta_n^* x_n^s]^{1/s}$ is *mean of order s* of the inputs, $x_1,...,x_N$ and they referred to Hardy, Littlewood and Polya (1934; 13) for the mathematical properties of this class of means.

18. Show that the CES production function f(x) defined by (61) above is homogeneous of degree one in the components of x.

19. Show that the CES production function f(x) defined by (61) above is a concave function of x if $s \neq 0$ and $s \leq 1$ and is a convex function of x if $s \geq 1$. *Hint*: Calculate the matrix of second order partial derivatives of f, $\nabla_{xx}^2 f(x)$, for $x >> 0_N$ and show it is negative semidefinite if $s \leq 1$ and positive semidefinite if $s \geq 1$. Problem 17 will be useful.

We now want to determine the cost minimizing system of input demand functions. We will first calculate the unit cost function that corresponds to the CES production function defined by (61). We assume that the producer faces the positive input prices $p \equiv [p_1,...,p_N] >> 0_N$. The unit cost minimization problem is the following one:

(62) $\min_x \{\Sigma_{n=1}^N p_n x_n : [\Sigma_{n=1}^N \beta_n x_n^s]^{1/s} = 1; x \geq 0_N\}$.

Ignoring the nonnegativity constraints, $x \geq 0_N$ and assuming that $s < 1$ and $s \neq 0$, the Lagrangian first order conditions for an interior solution for (62) are equivalent to the following conditions:

(63) $p_n = \lambda\beta_n x_n^{s-1}$; $\qquad n = 1,...,N;$
(64) $1 = \Sigma_{n=1}^N \beta_n x_n^s$

where the unknowns in (63) and (64) are $x_1,...,x_N$ and the Lagrange multiplier λ. The solution to (63) and (64) turns out to be the following one (remember, $s \neq 0$ and $s \neq 1$):[38]

(65) $x_n^*(p) \equiv p_n^{1/(s-1)}\beta_n^{1/(1-s)}/[\Sigma_{i=1}^N \beta_i^{1/(1-s)}p_n^{s/(s-1)}]^{1/s}$; $\qquad n = 1,...,N.$

Once the unit output demand functions have been calculated, the unit cost function, c(p), can be calculated:

(66) $c(p) \equiv \Sigma_{n=1}^N p_n x_n^*(p)$
$= [\Sigma_{n=1}^N \beta_n^{1/(1-s)}p_n^{s/(s-1)}]^{(s-1)/s}$ $\qquad$ using (65)
$= [\Sigma_{n=1}^N \alpha_n p_n^r]^{1/r}$

where the new parameters r and $\alpha_1, ..., \alpha_N$ are defined as follows:[39]

(67) $r \equiv s/(s-1)$; $\alpha_n \equiv \beta_n^{1/(1-s)}$; $n = 1,...,N.$

When s takes on the values between 1 and $-\infty$, $r = s/(s-1)$ goes from $-\infty$ to 1.[40] Thus the range of r and s is the same, but they travel in opposite directions. Hence the CES unit

---

[38] When s = 1, we have a linear production function. Usually, an interior solution to the cost minimization problem defined by (62) will not occur; i.e., in this case, we have a linear programming problem and the solution will normally be a corner solution.

[39] Note that $c(p) \equiv [\Sigma_{n=1}^N \alpha_n p_n^r]^{1/r}$ can be rewritten as $\gamma^*[\Sigma_{n=1}^N \alpha_n^* p_n^r]^{1/r}$ where $\alpha_n^* \equiv \alpha_n/\Sigma_{i=1}^N \alpha_i$ and $\gamma^* \equiv [\Sigma_{i=1}^N \alpha_i]^{1/r}$. Thus c(p) is equal to an efficiency parameter γ^* times a mean of order r.

cost function c(p) defined by (66) will be a linearly homogeneous, concave and nondecreasing function, and have the same mathematical properties as the CES production function f(x) defined by (61).

Once the CES unit cost function has been defined, the CES total cost function is defined as $C(y,p) \equiv yc(p)$ where c(p) is defined by (66). Using Shephard's Lemma, the CES system of cost minimizing demand functions is the following one:

$$(68)\ x_n(y,p) = y\,\alpha_n p_n^{r-1}[\Sigma_{i=1}^N \alpha_i\, p_i^r\,]^{(1/r)-1}\ ; \qquad n = 1,...,N.$$
$$= C(y,p)\,\alpha_n p_n^{r-1}/\Sigma_{i=1}^N \alpha_i\, p_i^r\ .$$

Problem

20. Recall problem 15 above which defined the Allen Uzawa *elasticity of substitution* σ_{nk} between inputs n and k. Show that if C(y,p) is the CES total cost function, then $\sigma_{nk}(y,p) = 1 - r$ for all input pairs n,k such that $n \neq k$. Thus every elasticity of substitution between any two distinct inputs is *equal to the same constant*.

The above problem shows why the CES functional form is unsatisfactory if the number of inputs N exceeds two, since it is a priori unlikely that all elasticities of substitution between every pair of inputs would equal the same number. Thus in the following sections, we will look for functional forms for the production or cost function that allow for more flexible patterns of substitution between inputs.

We conclude this section by listing some possible methods for estimating the elasticity of substitution if the underlying technology can be adequately described by the CES functional form.

We will first look at estimating equations where input prices are exogenous variables and input quantities (and hence output) are endogenous variables. Take logarithms of both sides of the CES input demand functions defined by (68). Add error terms to each equation, say e_n^t for equation n in period t.[41] Subtract the logarithm of the first input demand function from these N equations. Suppose that there are data on inputs, output, and input prices for t periods and the period t data are $x^t \equiv [x_1^t,...,x_N^t]$, y^t and $p^t \equiv [p_1^t,...,p_N^t]$ for t = 1,...,T. We obtain the following estimating equations:[42]

---

[40] Note that when s = 0, r will also equal 0. Rewrite the c(p) defined by the last line in (66) as $c(p) = \gamma^* c_r(p)$ where $\gamma^* \equiv [\Sigma_{i=1}^N \alpha_i]^{1/r}$ and $c_r(p) \equiv [\Sigma_{n=1}^N \alpha_n^* p_n^r\,]^{1/r}$. Using the results of problem 16, it can be seen that the limiting case for $c_r(p)$ as r tends to 0 is the Cobb-Douglas unit cost function which has the logarithm equal to $\Sigma_{n=1}^N \alpha_n^* \ln p_n$ where $\alpha_n^* = \alpha_n/\Sigma_{i=1}^N \alpha_i$ for n = 1,...,N.

[41] The errors in our models can be due to measurement errors in the prices and quantities, the assumption of incorrect functional forms and errors in optimization.

[42] Much of the literature on estimating CES unit cost functions deals with the application of this functional form in the consumer context when aggregating over similar products; e.g., see Broda and Weinstein (2010), Bernard, Redding and Schott (2010) and Gábór-Toth and Vermeulen (2017). Almost all of the estimating equations discussed in this section can be applied to the consumer context; i.e., replace the period t output level y^t by the period t utility level u^t and interpret x^t as a vector of cost minimizing

(69) $\ln[x_n^t/x_1^t] = \ln\alpha_n - \ln\alpha_1 + (r-1)\ln[p_n^t/p_1^t] + e_n^t - e_1^t$; n = 2,...,N; t = 1,...,T.

The above equations are linear in the unknown parameters, the $\ln\alpha_n$ and $r-1 \equiv -\sigma$. However, not all of the $\ln\alpha_n$ can be identified. This may not matter if the focus is on the estimation of r (or on the elasticity of substitution, σ). In order to identify all of the parameters, we can add a unit cost function equation to the system defined by (69) Thus define observed unit cost in period t as $c^t \equiv (\Sigma_{n=1}^N p_n^t x_n^t)/y^t$ for t = 1,...,T. Add the following estimating equations to equations (69) where e_0^t is the period t error term:[43]

(70) $\ln c^t = (1/r)\ln[\Sigma_{n=1}^N \alpha_n p_n^r] + e_0^t$; t = 1,...,T.

Of course, the estimating equations in (70) are nonlinear in the unknown parameters so nonlinear regression techniques will have to be used.

If the focus is on estimating the elasticity of substitution, equations (69) can be differenced again, this time with respect to time. Thus define the *double differenced logarithmic input quantity and price variables*, dx_n^t and dp_n^t as follows for n = 2,...,N; t = 2,...,T:

(71) $dx_n^t \equiv \ln[x_n^t/x_1^t] - \ln[x_n^{t-1}/x_1^{t-1}] = \ln x_n^t - \ln x_1^t - \ln x_n^{t-1} + \ln x_1^{t-1}$;
(72) $dp_n^t \equiv \ln[p_n^t/p_1^t] - \ln[p_n^{t-1}/p_1^{t-1}] = \ln p_n^t - \ln p_1^t - \ln p_n^{t-1} + \ln p_1^{t-1}$.

The double differenced counterparts to equations (69) are now the following equations:[44]

(73) $dx_n^t = (r-1)dp_n^t + e_n^t - e_1^t - e_n^{t-1} + e_1^{t-1}$; n = 2,...,N; t = 2,...,T

where $r-1 = -\sigma$. There are (N−1)(T−1) estimating equations in the system of equations defined by (73) and only one economic parameter to estimate, namely $-\sigma = r-1$. Note that the only exogenous variables in equations (69), (70) and (73) are input prices. Thus to prevent biased estimates, it is important that these prices be measured with minimal measurement error.

---

consumer demands. Estimating equations which involve y^t cannot be used in the consumer context since the utility level u^t is not observable.

[43] This equation cannot be estimated in the consumer context because unit cost c^t is not observable.

[44] The double differencing methodology originated in Feenstra (1994; 163). Equations (73) can be converted into double differenced log input shares equal to a constant times double differenced log input prices plus error terms; see Broda and Weinstein (2006; 564) (2010; 714) and Gábór-Toth and Vermeulen (2017) and equations (75) below for these share equations. The present analysis follows the material in Diewert and Feenstra (2017; 14). Diewert and Feenstra (2017; 76-79) worked out the analogous estimating equations for a CES direct aggregator function where double differenced log shares were equal to double differenced log quantities plus error terms. A potential cost of the double differencing technique is that the variance of the error terms in the system of estimating equations (73) can be much larger than the variances in the system of equations defined by equations (69) or in a system that just used x_n^t or $\ln x_n^t$ as the dependent variable for input n in period t. However, the standard error for σ when the very simple estimating system of equations defined by (76) used by Diewert and Feenstra was very small.

There is a problem with the systems of estimating equations defined by (69) and (73) and that is that these equations are dependent on the choice of the numeraire input, which in the above algebra is input 1. Looking at the estimating equations, it is evident that it is probably best to choose the numeraire commodity as one where the original error terms, the e_n^t, have means close to 0 and small variances. In practice, it may be difficult to choose the "best" numeraire commodity.[45]

There is a way to avoid asymmetry in the estimating equations and that is to shift from estimating systems of input demand functions to estimating systems of share equations. From equations (68), it can be seen that $s_n(y,p) \equiv p_n x_n(y,p)/C(y,p) = \alpha_n p_n^r/\Sigma_{i=1}^N \alpha_i p_i^r$ for n = 1,...,N. Define the nth input share of cost in period t as $s_n^t \equiv p_n^t x_n^t/\Sigma_{i=1}^N p_i^t x_i^t$ for n = 1,...,N and t =1,...,T. Adding error terms to the above cost share equations leads to the following nonlinear system of estimating equations:

$$(74)\ s_n^t = [\alpha_n(p_n^t)^r/\Sigma_{i=1}^N \alpha_i(p_i^t)^r] + e_n^t\ ; \qquad n = 1,...,N;\ t = 1,....,T.$$

If we sum equations (74) over n for a fixed t, we find that $\Sigma_{n=1}^N e_n^t = 0$ for t = 1,...,T. Thus within each time period, the errors cannot be distributed independently. Thus to prevent exact collinearity, one of the N estimating equations must be dropped. Furthermore, it can be seen that not all of the α_n parameters can be identified. Thus we require a normalization on the α_n such as $\Sigma_{i=1}^N \alpha_i = 1$ or $\alpha_1 = 1$. Alternatively, equations (70) can be added to the (N−1)T independent estimating equations in (74) as additional estimating equations which will enable all of the α_n to be identified.

An alternative stochastic specification can be obtained if we take logarithms of both sides of the equations $s_n^t = [\alpha_n(p_n^t)^r/\Sigma_{i=1}^N \alpha_i(p_i^t)^r]$ and add error terms e_n^{t*} to the resulting equations. Choose input 1 as a numeraire input and consider the following estimating equations:

$$(75)\ \ln(s_n^t/s_1^t) = \ln\alpha_n - \ln\alpha_1 + r\ln[p_n^t/p_1^t] + e_n^{t*} - e_1^{t*}\ ; \qquad n = 2,...,N;\ t = 1,...,T.$$

If the focus is on estimating the elasticity of substitution, $\sigma = 1 - r$, then equations (75) can be differenced with respect to time and we obtain the following system of estimating equations:

$$(76)\ ds_n^t = rdp_n^t + e_n^{t*} - e_1^{t*} - e_n^{t-1*} + e_1^{t-1*}\ ; \qquad n = 2,...,N;\ t = 2,...,T$$

[45] Here is a possible strategy for choosing the numeraire input. Take logs of both sides of equations (68) and add the error term (with 0 mean) e_n^t to equation n for period t. Run a preliminary systems nonlinear regression in order to obtain estimates for the variance-covariance matrix Σ of the vector of errors, $[e_1^t,...,e_N^t]$ which is assumed to be distributed independently over time. Use the estimated variance-covariance matrix, Σ^* say, to solve the convex programming problem, $\min_w \{w^T\Sigma^*w : w^Tw = 1\}$. The w^* solution to this problem will be a normalized eigenvector that corresponds to the smallest eigenvalue of Σ^*. Make a further normalization of w^* so that the resulting vector, w^{**}, satisfies the constraint $w^{**T}1_N = 1$. If Σ^* is a diagonal matrix, this methodology will pick the numeraire input to be the input which has the smallest error variance.

where the double differenced log price dp_n^t is defined by (72) and the double differenced log share ds_n^t is defined as $\ln s_n^t - \ln s_1^t - \ln s_n^{t-1} + \ln s_1^{t-1}$. Note that dp_n^t appears as an exogenous variable on the right hand sides of equation n,t in (73) and (76).

We conclude this section by considering the estimation of a system of CES inverse demand functions; i.e., we assume that prices are the endogenous variables and output and input quantities are the exogenous variables. Thus the input prices are regarded as the prices that rationalize the observed choice of inputs, assuming that the CES production function is the "true" production function.[46] This may seem to be an odd thing to do but it can turn out that estimating the CES system of inverse demand functions can lead to a much better fitting model than estimating the CES system of direct input demand functions as was done above.[47]

Let $y > 0$ and $p >> 0_N$ and the technology can be described by the CES production function defined by (61); i.e., $f(x_1,...,x_N) \equiv [\Sigma_{n=1}^N \beta_n x_n^s]^{1/s}$ where $s < 1$, $s \neq 0$ and $\beta_n > 0$ for $n = 1,...,N$. Then the producer's cost minimization problem is equivalent to the following constrained maximization problem:

(77) $\min_x \{\Sigma_{n=1}^N p_n x_n : \Sigma_{n=1}^N \beta_n x_n^s = y^s; x \geq 0_N\}$.

The first order necessary (and sufficient) conditions for solving (77) are equivalent to the following conditions:

(78) $p_n = \lambda\beta_n x_n^{s-1}$; $\qquad n = 1,...,N;$
(79) $y^s = \Sigma_{n=1}^N \beta_n x_n^s$.

Multiply both sides of equation n in equations (78) by x_n and sum the resulting equations. We obtain the following equation:

(80) $\Sigma_{n=1}^N p_n x_n = \lambda\Sigma_{n=1}^N \beta_n x_n^s$
$= \lambda y^s$

where the second equation follows using (79). Use the second equation in (80) to solve for $\lambda = \Sigma_{n=1}^N p_n x_n/y^s$ and substitute this equation back into equations (78). The resulting equations evaluated at the period t data are equations (81) below. As usual, the period t data are $x^t \equiv [x_1^t,...,x_N^t]$, y^t and $p^t \equiv [p_1^t,...,p_N^t]$ for $t = 1,...,T$. T We obtain the following equations:

(81) $p_n^t/(\Sigma_{n=1}^N p_n^t x_n^t) = \beta_n (x_n^t)^{s-1}/(y^t)^s$; $\qquad n = 1,...,N; t = 1,...,T.$

[46] This was the methodological approach taken by Arrow, Chenery, Minhas and Solow (1961) in their pioneering study on the estimation of CES functional forms. If the CES unit cost function model fits the observed data perfectly, then it will turn out that estimating the direct CES production function using a system of inverse demand functions will also fit the data perfectly.
[47] This was the case in the empirical study of CES estimation undertaken by Diewert and Feenstra (2017).

Take logarithms of both sides of equations (81) and add the error term e_n^t to the resulting equations.[48] We obtain the following system of estimating equations:

(82) $\ln[p_n^t/(\Sigma_{n=1}^N p_n^t x_n^t)] = \ln\beta_n + (s-1)\ln x_n^t - s\ln y^t + e_n^t$; $\quad n = 1,...,N;\ t = 1,...,T.$

Choose input 1 as the numeraire input and form the differenced equations (83):

(83) $\ln[p_n^t/p_1^t] = \ln\beta_n - \ln\beta_1 + (s-1)\ln[x_n^t/x_1^t] + e_n^t - e_1^t$; $\quad n = 2,3,...,N;\ t = 1,...,T.$

Not all of the parameters β_n can be identified using the (N−1)T equations in (83). In order to identify all of the β_n, we could make y an endogenous variable that is explained by the exogenous x_n, using the production function, $y = [\Sigma_{n=1}^N \beta_n x_n^s]^{1/s}$. Thus we could add the following estimating equations (84) to equations (83):

(84) $\ln y^t = (1/s)\ln[\Sigma_{n=1}^N \beta_n (x_n^t)^s] + e_0^t$; $\quad t = 1,...,T.$

If the focus is on estimating the elasticity of substitution, then we can time difference equations (83) and obtain the following estimating equations:

(85) $dp_n^t = (s-1)dx_n^t + e_n^t - e_1^t - e_n^{t-1} + e_1^{t-1}$; $\quad n = 2,...,N;\ t = 2,...,T$

where the double log differenced variables dx_n^t and dp_n^t are defined by (71) and (72). Recall the r which appeared in the CES cost function. The elasticity of substitution that corresponds to r is $\sigma = 1-r$. The s which appears in equations (85) corresponds to $r = s/(s-1)$. Thus $s-1 = -\sigma^{-1}$. Our previous system of estimating equations (73) for r can be written as $dx_n^t = (r-1)dp_n^t = -\sigma dp_n^t$, where we have omitted the error terms. Our new system of estimating equations for s, equations (85), can be written as $dp_n^t = (s-1)dx_n^t = -\sigma^{-1}dx_n^t$ where we have again omitted the error terms. Thus if either CES model fits the data perfectly, then the other model will fit the data perfectly and the two estimates for σ will be identical. Note that the two systems of estimating equations both have (N−1)(T−1) degrees of freedom and only one (non-variance) parameter, σ, to estimate.

It is useful to obtain a different system of estimating equations. Recall the first order condition equations (79) above. If we evaluate these equations using the period t data, we obtain the following equations which will hold if there are no errors in the CES cost minimization model:

(86) $(y^t)^s = \Sigma_{n=1}^N \beta_n (x_n^t)^s$; $\quad t = 1,...,T.$

Recall our earlier first order condition equations (81). Multiply equation n,t by x_n^t and we obtain the following system of equations after adding error terms, e_n^t:

(87) $s_n^t \equiv p_n^t x_n^t/(\Sigma_{i=1}^N p_i^t x_i^t) + e_n^t \quad n = 1,...,N;\ t = 1,...,T$
$\qquad = \beta_n (x_n^t)^s/(y^t)^s + e_n^t \quad$ using equations (81)

[48] These error terms are different from the error terms defined previously.

$= \beta_n (x_n^t)^s/\Sigma_{i=1}^N \beta_i(x_i^t)^s + e_n^t$; using equations (86).

If we sum equations (87) over n for a fixed t, we find that $\Sigma_{n=1}^N e_n^t = 0$ for t = 1,...,T. Thus within each time period, the errors cannot be distributed independently. To prevent exact collinearity, one of the N estimating equations must be dropped. Furthermore, it can be seen that not all of the β_n parameters can be identified. Thus we require a normalization on the β_n such as $\Sigma_{i=1}^N \beta_i = 1$ or $\beta_1 = 1$. Alternatively, equations (84) can be added to the (N−1)T independent estimating equations in (87) as additional estimating equations which will enable all of the β_n to be identified.

Note that the dependent variables in equations (87) are exactly the same as the dependent variables in our earlier nonlinear system of share estimating equations, equations (74). In equations (87), input quantities x_n^t are the explanatory variables whereas in equations (74), input prices p_n^t were the explanatory variables. In actual empirical applications of the CES model, the fit in the two systems can differ enormously.[49] This explains why we developed the algebra for the estimation of either system.

An alternative stochastic specification can be obtained if we take logarithms of both sides of the equations $s_n^t = [\beta_n(p_n^t)^r/\Sigma_{i=1}^N \beta_i(p_i^t)^r]$ and add error terms e_n^{t*} to the resulting equations. Choose input 1 as a numeraire input and consider the following estimating equations:

(88) $\ln(s_n^t/s_1^t) = \ln\beta_n - \ln\beta_1 + s\ln[x_n^t/x_1^t] + e_n^{t*} - e_1^{t*}$; n = 2,...,N; t = 1,...,T.

If the focus is on estimating the elasticity of substitution, $\sigma = 1/(1 - s)$, then equations (88) can be differenced with respect to time and we obtain the following system of estimating equations:

(89) $ds_n^t = sdx_n^t + e_n^{t*} - e_1^{t*} - e_n^{t-1*} + e_1^{t-1*}$; n = 2,...,N; t = 2,...,T

where the double differenced log input quantity dx_n^t is defined by (73) and the double differenced log share ds_n^t is defined as $\ln s_n^t - \ln s_1^t - \ln s_n^{t-1} + \ln s_1^{t-1}$. Note that dx_n^t appears as an exogenous variable on the right hand sides of equation n,t in (85) and (89).

8. Flexible Functional Forms for Cost Functions: The Generalized Leontief Functional Form

From the previous section, it can be seen that the CES functional form is not suitable for economic applications where elasticities of substitution are allowed to be different between different pairs of inputs. This leads us to define formally the concept of a *flexible functional form*. We will define this concept first for a unit cost function c(p) and then for a general cost function C(y,p).

[49] See Diewert and Feenstra (2017). The system (87) fit their data much better than the corresponding system (74).

Let $c^*(p)$ be an arbitrary unit cost function that satisfies the appropriate regularity conditions on unit cost functions and in addition, is twice continuously differentiable around a point $p^* >> 0_N$. Then we say that a unit cost function c(p) that is also twice continuously differentiable around the point p^* is *flexible* if it has enough free parameters so that the following $1 + N + N^2$ equations can be satisfied:[50]

(90) $c(p^*) = c^*(p^*)$;
(91) $\nabla c(p^*) = \nabla c^*(p^*)$;
(92) $\nabla^2 c(p^*) = \nabla^2 c^*(p^*)$.

Thus c(p) is a flexible functional form if it has enough free parameters to provide a second order Taylor series approximation to an arbitrary unit cost function.

At first glance, it looks like c(p) will have to have at least $1 + N + N^2$ independent parameters in order to be able to satisfy all of the equations (90)-(92). However, since both c and c^* are assumed to be twice continuously differentiable, Young's Theorem in calculus implies that $\partial^2 c(p^*)/\partial p_i \partial p_k = \partial^2 c(p^*)/\partial p_k \partial p_i$ for all $i \neq k$ (and of course, the same equations hold for the second order partial derivatives of $c^*(p)$ when evaluated at $p = p^*$). Thus the N^2 equations in (92) can be replaced with the following N(N+1)/2 equations:

(93) $\partial^2 c(p^*)/\partial p_i \partial p_k = \partial^2 c^*(p^*)/\partial p_i \partial p_k$; for $1 \leq i \leq k \leq N$.

Another property that both unit cost functions must have is homogeneity of degree one in the components of p. By part 1 of Euler's Theorem on homogeneous functions, c and c^* satisfy the following equations:

(94) $c(p^*) = p^{*T}\nabla c(p^*)$ and $c^*(p^*) = p^{*T}\nabla c^*(p^*)$.

Thus if c and c^* satisfy equations (91), then using (94), we see that c and c^* automatically satisfy equation (90). By part 2 of Euler's Theorem on homogeneous functions, c and c^* satisfy the following equations:

(95) $\nabla^2 c(p^*)p^* = 0_N$ and $\nabla^2 c^*(p^*)p^* = 0_N$.

This means that if we have $\partial^2 c(p^*)/\partial p_i \partial p_k = \partial^2 c^*(p^*)/\partial p_i \partial p_k$ for all $i \neq k$, then equations (95) will imply that $\partial^2 c(p^*)/\partial p_i \partial p_i = \partial^2 c^*(p^*)/\partial p_i \partial p_i$ as well, for i = 1,…,N.

Summarizing the above material, if c(p) is linearly homogeneous, then in order for it to be flexible, c(p) needs to have only enough parameters so that the N equations in (91) can be satisfied and so that the following N(N–1)/2 equations can be satisfied:

(96) $\partial^2 c(p^*)/\partial p_j \partial p_k = \partial^2 c^*(p^*)/\partial p_i \partial p_k \equiv c_{ik}^*$; for $1 \leq i < k \leq N$.

[50] Diewert (1971) introduced the concept of a flexible functional form. The actual term "flexible" was introduced in Diewert (1974a; 133).

Thus in order to be flexible, c(p) must have at least $N + N(N-1)/2 = N(N+1)/2$ independent parameters.

Recall that the Generalized Leontief cost function was introduced in Section 4. The unit cost function that corresponds to this function form is defined as follows:[51]

$$(97)\ c(p) \equiv \sum_{i=1}^{N}\sum_{k=1}^{N} b_{ik}\, p_i^{1/2}\, p_k^{1/2}$$

where $b_{ik} = b_{ki}$ for all i and k. Note that there are exactly N(N+1)/2 independent b_{ik} parameters in the c(p) defined by (97). For this functional form, the N equations in (91) become:

$$(98)\ \partial c(p^*)/\partial p_n = \sum_{k=1}^{N} b_{nk}\, (p_k^*/p_n^*)^{1/2} = \partial c^*(p^*)/\partial p_n \equiv c_n^* ; \qquad n = 1,\ldots,N.$$

The N(N−1)/2 equations in (96) become:

$$(99)\ (½)b_{ik} /(p_i^* p_k^*)^{1/2} = c_{ik}^* ; \qquad 1 \le i < k \le N.$$

However, it is easy to solve equations (99) for the b_{ik}:

$$(100)\ b_{ik} = 2c_{ik}^*(p_i^* p_k^*)^{1/2} ; \qquad 1 \le i < k \le N.$$

Once the b_{ik} for $i < k$ have been determined using (100), we set $b_{ki} = b_{ik}$ for $i < k$ and finally the b_{ii} are determined using the N equations in (98).

The above material shows how we can find a flexible functional form for a unit cost function[52]. We now turn our attention to finding a flexible functional form for a general cost function C(y,p). Let $C^*(y,p)$ be an arbitrary cost function that satisfies the appropriate regularity conditions on cost functions listed in Theorem 1 above and in addition, is twice continuously differentiable around a point (y^*,p^*) where $y^* > 0$ and $p^* >> 0_N$. Then we say that a given cost function C(y,p) that is also twice continuously differentiable around the point (y^*,p^*) is *flexible* if it has enough free parameters so that the following $1 + (N+1) + (N+1)^2$ equations can be satisfied:

$$(101)\quad C(y^*,p^*) = C^*(y^*,p^*) ; \qquad \text{(1 equation)}$$
$$(102)\quad \nabla_p C(y^*,p^*) = \nabla_p C^*(y^*,p^*) ; \qquad \text{(N equations)}$$
$$(103)\quad \nabla^2_{pp} C(y^*,p^*) = \nabla^2_{pp} C^*(y^*,p^*) ; \qquad \text{(}N^2\text{ equations)}$$
$$(104)\quad \nabla_y C(y^*,p^*) = \nabla_y C^*(y^*,p^*) ; \qquad \text{(1 equation)}$$
$$(105)\quad \nabla^2_{py} C(y^*,p^*) = \nabla^2_{py} C^*(y^*,p^*) ; \qquad \text{(N equations)}$$
$$(106)\quad \nabla^2_{yp} C(y^*,p^*) = \nabla^2_{yp} C^*(y^*,p^*) ; \qquad \text{(N equations)}$$
$$(107)\quad \nabla^2_{yy} C(y^*,p^*) = \nabla^2_{yy} C^*(y^*,p^*) ; \qquad \text{(1 equation).}$$

---

[51] We no longer restrict the b_{ij} to be nonnegative.

[52] This material can be adapted to the case where we want a flexible functional form for a linearly homogeneous utility or production function f(x): just replace p by x and c(p) by f(x).

Equations (101)-(107) are the counterparts to our earlier unit cost equations (90)-(92). As was the case with unit cost functions, equation (102) is implied by the linear homogeneity in prices of the cost functions and Part 1 of Euler's Theorem on homogeneous functions. Young's Theorem on the symmetry of cross partial derivatives means that the lower triangle of equations in (103) is implied by the equalities in the upper triangle of both matrices of partial derivatives. Part 2 of Euler's Theorem on homogeneous functions implies that if all the off diagonal elements in both matrices in (103) are equal, then so are the diagonal elements. Hence, in order to satisfy all of the equations in (101)-(103), we need only satisfy the N equations in (102) and the N(N−1)/2 equations in the upper triangle of the N^2 equations in (103). Young's Theorem implies that if equations (105) are satisfied, then so are equations (106). However, Euler's Theorem on homogeneous functions implies that

(108) $\partial C(u^*,p^*)/\partial y = p^{*T}\nabla^2_{py}C(y^*,p^*) = p^{*T}\nabla^2_{py}C^*(u^*,p^*) = \partial C^*(y^*,p^*)/\partial y$.

Hence, if equations (105) are satisfied, then so is the single equation (104). Putting this all together, we see that in order for C to be flexible, we need enough free parameters in C so that the following equations can be satisfied:

- Equations (102); N equations;
- The upper triangle in equations (103); N(N−1)/2 equations;
- Equations (105); N equations; and
- Equation (107); 1 equation.

Hence, in order for C to be a flexible functional form, it will require a minimum of 2N + N(N−1)/2 +1 = N(N+1)/2 + N +1 parameters. Thus a fully flexible cost function, C(y,p), will require N + 1 additional parameters compared to a flexible unit cost function, c(p).

In the following Sections, we will define several flexible functional forms for unit cost functions c(p). Once we have a flexible functional form for a unit cost function c(p), then the algebra below shows how we can modify c(p) to obtain a flexible total cost function C(y,p).[53]

Suppose the unit cost function is the Generalized Leontief unit cost function c(p) defined by (97) above. We now show how terms can be added to it in order to make it a fully flexible cost function. Thus define C(u,p) as follows:

(109) $C(y,p) \equiv yc(p) + b^Tp + (1/2)a_0\alpha^Tpy^2$

where $b \equiv [b_1,\ldots,b_N]$ is an N dimensional vector of new parameters, a_0 is a new parameter and $\alpha \equiv [\alpha_1,\ldots,\alpha_N] > 0_N$ is a vector of predetermined parameters.[54] Using (109) as our

[53] The algebra for converting the translog unit cost function into the translog cost function is different.

[54] We have defined the cost function C in this manner so that it has the *minimal* number of parameters required in order to be a flexible functional form. Thus it is a *parsimonious* flexible functional form.

candidate for a flexible (total) cost function C, equations (102), (103), (105) and (107) become:

$$(110)\quad y^*\nabla_p c(p^*) + b + (1/2)a_0\alpha y^{*2} = \nabla_p C^*(y^*,p^*) ;$$
$$(111)\quad y^*\nabla^2_{pp} c(p^*) = \nabla^2_{pp} C^*(y^*,p^*) ;$$
$$(112)\quad \nabla_p c(p^*) + a_0\alpha y^* = \nabla^2_{py} C^*(y^*,p^*) ;$$
$$(113)\quad a_0\alpha^T p^* = \nabla^2_{yy} C^*(y^*,p^*) .$$

Use equations (111) in order to determine the b_{ik} for $i \neq k$. Use (113) in order to determine the single parameter a_0. Use equations (112) in order to determine the b_{ii}. Finally, use equations (110) in order to determine the parameters b_n in the b vector. Thus the cost function C(u,p) defined by (109), which uses the Generalized Leontief unit cost function c(p) defined by (97) as a building block, is a parsimonious flexible functional form for a general cost function.

In fact, it is not necessary to use the Generalized Leontief unit cost function in definition (109) in order to convert a flexible functional form for a unit cost function into a flexible functional form for a general cost function. Let c(p) be any flexible functional form for a unit cost function and define C(y,p) by (109). Use equation (113) to determine the parameter a_0. Once a_0 has been determined, equations (111) and (112) can be used to determine the parameters in the unit cost function c(p). Finally, equations (110) can be used to determine the parameters in the vector b.

Differentiating (109) leads to the following system of estimating equations, where $x(y,p) = \nabla_p C(y,p)$ is the producer's system of cost minimizing input demand functions:

$$(114)\quad x(y,p) = y\nabla c(p) + b + (1/2)a_0\alpha y^2.$$

If the Generalized Leontief unit cost function is used as the c(p) in equations (114), then the N estimating equations will be linear in the unknown parameters. This will facilitate econometric estimation. The cross equation symmetry restrictions could be tested or imposed.

In empirical applications, if we use the Generalized Leontief functional form when there are more than two inputs, a problem can occur: one or more of the estimated b_{ik} can turn out to be negative numbers (so that inputs i and k are complements). Under these conditions, the estimated cost function can fail to be concave at the observed data points and it will not be globally concave over all positive input prices. Global concavity can be imposed by replacing the off diagonal b_{ik} parameters by their squares[55] but if this is done, then all pairs of inputs will be either substitutes or be unrelated. Global concavity can be imposed but at the cost of destroying the flexibility of the functional form.[56] Thus the Generalized Leontief functional form is not a "perfect" flexible functional form. Finding

---

[55] The resulting estimating equations become nonlinear in the parameters when we square the b_{ik}. Typically, this does not create any problems: just use a nonlinear estimation method.

[56] If there are more than 4 inputs and we allow for complementarity, then experience has shown that complementary input pairs show up almost always.

flexible functional forms where the restrictions implied by microeconomic theory can be *imposed* on the functional form without destroying its flexibility is a nontrivial task which we will address later in Sections 10 and 11 below.

9. The Translog Functional Form

The translog unit cost function, c(p), is defined as follows:[57]

$$(115)\ \ln c(p) \equiv \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik} \ln p_i \ln p_k$$

where the parameters α_i and γ_{ik} satisfy the following restrictions:

$$(116)\ \gamma_{ik} = \gamma_{ki}\ ; \quad 1 \le i < k \le N\ ; \quad \text{(N(N−1)/2 symmetry restrictions)}$$
$$(117)\ \sum_{i=1}^{N} \alpha_i = 1\ ; \quad \text{(1 restriction)}$$
$$(118)\ \sum_{k=1}^{N} \gamma_{ik} = 0\ ; \quad i = 1,...,N \quad \text{(N restrictions).}$$

Note that the symmetry restrictions (116) and the restrictions (118) imply the following restrictions:

$$(119)\ \sum_{i=1}^{N} \gamma_{ik} = 0\ ; \qquad k = 1,...,N.$$

There are 1+N α_i parameters and N^2 γ_{ik} parameters. However, the restrictions (116)-(119) mean that there are only N independent α_i parameters and N(N−1)/2 independent γ_{ik} parameters, which is the minimal number of parameters required for a unit cost function to be flexible.

We show that the translog unit cost function c(p) defined by (115)-(118) is linearly homogeneous; i.e., we need to show that $c(\lambda p) = \lambda c(p)$ for $\lambda > 0$ and $p >> 0_N$. Thus, we need to show that

$$(120)\ \ln c(\lambda p) = \ln[\lambda c(p)] = \ln\lambda + \ln c(p)\ ; \qquad \lambda > 0 \text{ and } p >> 0_N.$$

Using definition (115), we have

$$\begin{aligned}
(121)\ \ln c(\lambda p_1,..., \lambda p_N) &= \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln\lambda p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik} \ln\lambda p_i \ln\lambda p_k \\
&= \alpha_0 + \sum_{i=1}^{N} \alpha_i[\ln\lambda+\ln p_i] + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln\lambda+\ln p_i][\ln\lambda+\ln p_k] \\
&= \alpha_0 + \sum_{i=1}^{N} \alpha_i[\ln\lambda]+\sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln\lambda+\ln p_i][\ln\lambda+\ln p_k] \\
&= \alpha_0 + 1\,[\ln\lambda]+\sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln\lambda+\ln p_i][\ln\lambda+\ln p_k] \text{ using (117)} \\
&= \ln\lambda + \alpha_0+\sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln\lambda][\ln\lambda] \\
&\quad + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln\lambda][\ln p_k]+ (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln p_i][\ln\lambda] \\
&\quad + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln p_i][\ln p_k] \\
&= \ln\lambda + \alpha_0+\sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} [\sum_{k=1}^{N} \gamma_{ik}][\ln\lambda][\ln\lambda] \\
&\quad + (1/2) \sum_{k=1}^{N} [\sum_{i=1}^{N} \gamma_{ik}][\ln p_k][\ln\lambda] + (1/2) \sum_{i=1}^{N} [\sum_{k=1}^{N} \gamma_{ik}][\ln p_i][\ln\lambda]
\end{aligned}$$

[57] This functional form is due to Christensen, Jorgenson and Lau (1971) (1973) (1975). The material in this Section is due to these authors.

$$+ (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln p_i][\ln p_k]$$
$$= \ln\lambda + \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} [0][\ln\lambda][\ln\lambda]$$
$$+ (1/2) \sum_{k=1}^{N} [0][\ln p_k][\ln\lambda] + (1/2) \sum_{i=1}^{N} [0][\ln p_i][\ln\lambda]$$
$$+ (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik}[\ln p_i][\ln p_k] \qquad \text{using (118) and (119)}$$
$$= \ln\lambda + \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ij}[\ln p_i][\ln p_k]$$
$$= \ln\lambda + \ln c(p) \qquad \text{using definition (115)}$$

which establishes the linear homogeneity property (120). Thus the restrictions (116)-(118) imply the linear homogeneity of the translog unit cost function.

To establish the flexibility of the translog unit cost function c(p) defined by (115)-(118), we need only solve the following system of equations, which is equivalent to the N(N+1)/2 equations defined by (91) and (93):

(122) $\ln c(p) = \ln c^*(p^*)$; 1 equation
(123) $\partial \ln c(p^*)/\partial \ln p_i = \partial \ln c^*(p^*)/\partial \ln p_i$; $i = 1,2,...,N-1$; $N-1$ equations
(124) $\partial^2 \ln c(p^*)/\partial \ln p_i \partial \ln p_k = \partial^2 \ln c^*(p^*)/\partial \ln p_i \partial \ln p_k$; $1 \leq i < k \leq N$; $N(N-1)/2$ equations.

Upon differentiating the translog unit cost function defined by (115), we see that equations (123) are equivalent to the following equations:

(125) $\alpha_i + \sum_{k=1}^{N} \gamma_{ik} \ln p_j = \partial \ln c^*(p^*)/\partial \ln p_i$; $i = 1,2,...,N-1$.

Differentiating the translog unit cost function again, we find that equations (124) are equivalent to the following equations:

(126) $\gamma_{ik} = \partial^2 \ln c^*(p^*)/\partial \ln p_i \partial \ln p_k$; $1 \leq i < j \leq N$.

Now use equations (126) to determine the γ_{ik} for $1 \leq i < k \leq N$. Use the symmetry restrictions (116) to determine the γ_{ik} for $1 \leq k < i \leq N$. Use equations (118) to determine the γ_{ii} for $i = 1,2,...,N$. With the entire N by N matrix of the γ_{ij} now determined, use equations (125) in order to determine the α_i for $i = 1,2,...,N-1$. Now use equation (117) to determine α_N. Finally, use equation (112) to determine α_0.

We turn our attention to the problems involved in obtaining estimates for the unknown parameters α_i and γ_{ik}, which occur in the definition of the translog unit cost function, c(p) defined by (115). The total cost function C(y,p) is defined in terms of the unit cost function c(p) as follows:

(127) $C(y,p) \equiv yc(p)$.

Taking logarithms on both sides of (127) yields, after some rearrangement:

(128) $\ln[C(y,p)/y] = \ln c(p)$
$$= \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln p_i + (1/2) \sum_{i=1}^{N} \sum_{k=1}^{N} \gamma_{ik} \ln p_i \ln p_k$$

where we have replaced lnc(p) using (115). The corresponding system of cost minimizing input demand functions x(y,p) is obtained using Shephard's Lemma:

(129) $x(y,p) \equiv \nabla_p C(y,p) = y\nabla_p c(p)$.

Suppose that we have data for a production unit on output in period t, y^t, inputs $x^t \equiv [x_1^t,...,x_N^t]$ and input prices $p^t \equiv [p_1^t,...,p_N^t]$ for t = 1,...,T. Thus the *period t observed unit cost* is:

(130) $c^t \equiv p^{tT}x^t/y^t \equiv \sum_{i=1}^N p_i^t x_i^t/y^t$; t = 1,...,T.

Evaluate (128) at the period t data and add an error term, e_0^t. Using (130), (128) evaluated at the period t data becomes the following estimating equation:

(131) $\ln c^t = \alpha_0 + \sum_{i=1}^N \alpha_i \ln p_i^t + (1/2) \sum_{i=1}^N \sum_{k=1}^N \gamma_{ij} \ln p_i^t \ln p_k^t + e_0^t$; t = 1,...,T.

Note that (131) is linear in the unknown parameters.

In order to obtain additional estimating equations, we have to use the input demand functions, $x_i(y,p) \equiv y\partial c(p)/\partial p_i$ for i = 1,...,N; (see equations (129) above). The ith input share function, $s_i(y,p)$, is defined as:

(132) $s_i(y,p) \equiv p_i x_i(y,p)/C(y,p)$ i = 1,...,N
$= p_i[y\partial c(p)/\partial p_i]/C(y,p)$ using (129)
$= p_i[y\partial c(p)/\partial p_i]/yc(p)$ using (127)
$= p_i[\partial c(p)/\partial p_i]/c(p)$
$= \partial \ln c(p)/\partial \ln p_i$
$= \alpha_i + \sum_{k=1}^N \gamma_{ik} \ln p_k$

where the last equation follows upon differentiating the c(p) defined by (115).

Now evaluate both sides of (132) at the period t data and add error terms e_i^t to obtain the following system of estimating equations:

(133) $s_i^t \equiv p_i^t x_i^t/C^t = \alpha_i + \sum_{j=1}^N \gamma_{ij} \ln p_j^t + e_i^t$; i = 1,...,N; t = 1,...,T.

Note that equations (133) are also linear in the unknown parameters.[58] Obviously, the N estimating equations in (133) could be added to the single estimating equation (131) in order to obtain N+1 estimating equations with cross equation equality constraints on the parameters α_i and γ_{ij}. However, since total cost in any period t, C^t, equals the sum of the individual expenditures on the inputs, $\sum_{i=1}^N p_i^t x_i^t$, the observed input shares $s_i^t \equiv p_i^t x_i^t/C^t$ will satisfy the following constraint for each period t:

[58] Note also that the cross equation symmetry conditions, $\gamma_{ik} = \gamma_{ki}$, could be tested or imposed.

(134) $\sum_{i=1}^{N} s_i^t = 1$; $t = 1,...,T$.

Thus the stochastic error terms e_i^t in equations (133) cannot all be independent. Hence we must drop one estimating equation from (133). Thus equation (131) and any N−1 of the N equations in (133) may be used as a system of estimating equations in order to determine the parameters of the translog unit cost function.[59]

We now turn our attention to the problem of deriving a formula for the price elasticities of demand, $\partial x_i(y,p)/\partial p_j$, given that the unit cost function has the translog functional form defined by (115)-(118). Recall equations (132) above. For $k \neq i$, differentiate the ith equation in (132) with respect to the log of p_k and we obtain the following equations for all $k \neq i$:

(135) $\partial s_i(y,p)/\partial \ln p_k = p_i\{\partial[x_i(y,p)/C(y,p)]/\partial \ln p_k\} = \gamma_{ik}$.

Hence upon noting that $s_i(y,p) = p_i x_i(y,p)/C(y,p)$ and using (135), we have for $k \neq i$:

(136) $\gamma_{ik} = p_i\, \partial[x_i(y,p)/C(y,p)]/\partial \ln p_k$
$= p_i p_k \partial[x_i(y,p)/C(y,p)]/\partial p_k$
$= p_i p_k\{[1/C(y,p)][\partial x_i(y,p)/\partial p_k] - x_i(y,p)[1/C(y,p)]^2[\partial C(y,p)/\partial p_k]\}$
$= [p_i x_i(y,p)/C(y,p)]\{\partial \ln x_i(y,p)/\partial \ln p_k\} - [p_i x_i(y,p)/C(y,p)][p_k x_k(y,p)/C(y,p)]$
using Shephard's Lemma, $x_k(y,p) = \partial C(y,p)/\partial p_k$
$= s_i(y,p)\{\partial \ln x_i(y,p)/\partial \ln p_k\} - s_i(y,p)s_k(y,p)$.

Equations (136) can be rearranged to give us the following formula for the *cross price elasticities of input demand* for all $i \neq k$:

(137) $\partial \ln x_i(y,p)/\partial \ln p_k = [s_i(y,p)]^{-1}\gamma_{ik} + s_k(y,p)$.

Now differentiate the ith equation in (135) with respect to the logarithm of p_i and get the following equations:

(138) $\gamma_{ii} = p_i \partial[p_i x_i(y,p)/C(y,p)]/\partial p_i$; $i = 1,...,N$;
$= p_i[x_i(y,p)/C(y,p)] + [p_i/C(y,p)][\partial x_i(y,p)/\partial p_i] - [p_i x_i(y,p)/C(y,p)^2][\partial C(y,p)/\partial p_i]\}$
$= p_i\{[x_i(y,p)/C(y,p)] + [p_i/C(y,p)][\partial x_i(y,p)/\partial p_i] - [p_i x_i(y,p)/C(y,p)^2][x_i(y,p)]\}$

[59] In situations where N is large relative to the number of observations T, maximum likelihood estimation of equation (131) and N−1 of the equations (133) can fail if a general variance covariance matrix is estimated for the error terms in these equations. The problem is that all of the unknown economic parameters are contained in equation (131) and as a result, the estimated squared residuals in this equation will tend to be small relative to the estimated squared residuals in equations (133), where each equation has only a few unknown economic parameters. Hence equation (131) can suffer from multicollinearity problems and the small apparent variance of the residuals in this equation can lead to the maximum likelihood estimation procedure giving too much weight to the unit cost function equation relative to the other equations. Under these conditions, the resulting elasticities may be erratic and they may not satisfy the appropriate curvature conditions. Note that the estimation of the Generalized Leontief unit cost function did not suffer from this problem of having every unknown parameter in a single equation.

using Shephard's Lemma, $x_i(y,p) = \partial C(y,p)/\partial p_i$

$$= p_i x_i(y,p)/C(y,p) + [p_i x_i(y,p)/C(y,p)][\partial \ln x_i(y,p)/\partial \ln p_i] - [p_i x_i(y,p)/C(y,p)]^2$$
$$= s_i(y,p) + s_i(y,p)[\partial \ln x_i(y,p)/\partial \ln p_i] - s_i(y,p)^2.$$

Equations (138) can be rearranged to give us the following formula for the *own price elasticities of input demand*:

$$(139)\ \partial \ln x_i(y,p)/\partial \ln p_i = [s_i(y,p)]^{-1}\, \gamma_{ii} + s_i(y,p) - 1\ ; \qquad i = 1,\ldots,N.$$

Thus given econometric estimates for the α_i and γ_{ij}, which we denote by α_i^* and γ_{ij}^*, the estimated or fitted shares in period t, s_i^{t*} are defined using these estimates and equations (133) evaluated at the period t data:

$$(140)\ s_i^{t*} \equiv \alpha_i^* + \textstyle\sum_{j=1}^N \gamma_{ij}^* \ln p_j^t\ ; \qquad i = 1,\ldots,N\ ;\ t = 1,\ldots,T.$$

Now use equations (137) evaluated at the period t data and econometric estimates to obtain the following formula for the *period t cross elasticities of demand*, E_{ik}^t:

$$(141)\ E_{ik}^t \equiv \partial \ln x_i(y^t,p^t)/\partial \ln p_k = [s_i^{t*}]^{-1}\gamma_{ik}^* + s_k^{t*}\ ; \qquad i \neq k\ ;\ t = 1,\ldots,T.$$

Similarly, use equations (139) evaluated at the period t data and econometric estimates to obtain the following formula for the *period t own elasticities of demand*, E_{ii}^t:

$$(142)\ E_{ii}^t \equiv \partial \ln x_i(y^t,p^t)/\partial \ln p_i = [s_i^{t*}]^{-1}\gamma_{ii}^* + s_i^{t*} - 1\ ; \qquad i = 1,\ldots,N\ ;\ t = 1,\ldots,T.$$

We can also obtain an estimated or *fitted period t unit cost*, c^{t*}, by using our econometric estimates for the parameters and by exponentiating the right hand side of equation t in (130):

$$(143)\ c^{t*} \equiv \exp[\alpha_0^* + \textstyle\sum_{i=1}^N \alpha_i^* \ln p_i^t + (1/2) \sum_{i=1}^N \sum_{k=1}^N \gamma_{ik}^* \ln p_i^t \ln p_k^t]\ ; \qquad t = 1,\ldots,T.$$

Finally, our fitted period t shares s_i^{t*} defined by (50) and our fitted period t costs C^{t*} defined by (53) can be used in order to obtain estimated or *fitted period t input demands*, x_i^{t*}, as follows:

$$(144)\ x_i^{t*} \equiv y^t c^{t*} s_i^{t*}/p_i^t\ ; \qquad i = 1,\ldots,N\ ;\ t = 1,\ldots,T.$$

Given the matrix of period t estimated input price elasticities of demand, $[E_{ik}^t]$, we can readily calculate the matrix of period t *estimated input price derivatives*, $\nabla_p x(y^t,p^t) = \nabla^2_{pp}C(y^t,p^t) = y^t\nabla^2_{pp}c(p^t)$. The estimate for element ik of $\nabla^2_{pp}C(y^t,p^t)$ is:

$$(145)\ C_{ik}^{t*} \equiv E_{ik}^t\, x_i^{t*}/p_k^t\ ; \qquad i,k = 1,\ldots,N\ ;\ t = 1,\ldots,T$$

where the estimated period t elasticities E_{ik}^t are defined by (141) and (142) and the fitted period t input demands x_i^{t*} are defined by (144). Once the estimated input price derivative matrices $[C_{ik}^{t*}]$ have been calculated for period t, then we may check whether it is negative semidefinite using determinantal conditions or by checking if all of the eigenvalues of each matrix are zero or negative for t = 1,...,T. Unfortunately, *very frequently these negative semidefiniteness conditions will fail to be satisfied for both the translog and generalized Leontief functional forms*. Thus the translog and Generalized Leontief functional forms both suffer from the same problem: in general, it is not possible to impose concavity on these functional forms without destroying their flexibility property. Hence, in the following two sections, we study functional forms where these curvature conditions can be imposed without destroying the flexibility of the functional form.

10. The Normalized Quadratic Unit Cost Function.

The normalized quadratic unit cost function c(p) is defined as follows for $p >> 0_N$:[60]

(146) $c(p) \equiv b^T p + (1/2)p^T Bp/\alpha^T p$

where $b^T \equiv [b_1,...,b_N]$ and $\alpha^T \equiv [\alpha_1,...,\alpha_N]$ are parameter vectors and $B \equiv [b_{ik}]$ is a matrix of parameters. The vector α and the matrix B satisfy the following restrictions:

(147) $\alpha > 0_N$;
(148) $B = B^T$; i.e., the matrix B is symmetric;
(149) $Bp^* = 0_N$ for some $p^* >> 0_N$.

In most empirical applications, the vector of nonnegative but nonzero parameters α is fixed a priori. The two most frequent a priori choices for α are $\alpha \equiv 1_N$, a vector of ones or $\alpha \equiv (1/T)\sum_{t=1}^T x^t$, the sample mean of the observed input vectors. The two most frequent choices for the reference price vector p^* are $p^* \equiv 1_N$ or $p^* \equiv p^t$ for some period t; i.e., in this second choice, we simply set p^* equal to the observed period t price vector.

Assuming that α has been predetermined, there are N unknown parameters in the b vector and N(N−1)/2 unknown parameters in the B matrix, taking into account the symmetry restrictions (148) and the N linear restrictions in (149). Note that the c(p) defined by (146) is linearly homogeneous in the components of the input price vector p.

Another possible way of defining the normalized quadratic unit cost function is as follows:

(150) $c(p) \equiv (1/2)\, p^T Ap/\alpha^T p$

where the parameter matrix A is symmetric; i.e., $A = A^T \equiv [a_{ik}]$ and $\alpha > 0_N$ as before. Assuming that the vector of parameters α has been predetermined, the c(p) defined by (150) has N(N+1)/2 unknown a_{ik} parameters.

---

[60] This functional form was introduced by Diewert and Wales (1987; 53) where it was called the Symmetric Generalized McFadden functional form. It is a generalization of a functional form due to McFadden (1978; 279). Additional material on this functional form can be found in Diewert and Wales (1988) (1992) (1993).

Comparing (146) with (150), it can be seen that (150) has dropped the b vector but has also dropped the N linear constraints (149). It can be shown that the model defined by (146) is a special case of the model defined by (150). To show this, given (146), define the matrix A in terms of B, b and α as follows:

(151) $A \equiv B + [b\alpha^T + \alpha b^T]$.

Substituting (151) into (150), (150) becomes:

(152) $c(p) = (1/2)p^T\{B + [b\alpha^T + \alpha b^T]\}p/\alpha^T p$
$= (1/2)p^T Bp/\alpha^T p + (1/2)\, p^T[b\alpha^T + \alpha b^T]p/\alpha^T p$
$= (1/2)p^T Bp/\alpha^T p + (1/2)\{p^T b\alpha^T p + p^T\alpha b^T p\}/\alpha^T p$
$= (1/2)p^T Bp/\alpha^T p + (1/2)\{2p^T b\alpha^T p\}/\alpha^T p$
$= (1/2)p^T Bp/\alpha^T p + p^T b$

which is the same functional form as (146). However, it is preferable to work with the model (146) rather than with the seemingly more general model (150) for three reasons:

- The c(p) defined by (146) clearly contains the no substitution Leontief functional form as a special case (simply set $B = 0_{NxN}$);
- the estimating equations that correspond to (146) will contain constant terms and
- it is easier to establish the flexibility property for (146) than for (150).

The first and second order partial derivatives of the normalized quadratic unit cost function defined by (146) are given by:

(153) $\nabla_p c(p) = b + (\alpha^T p)^{-1}Bp - (1/2)(\alpha^T p)^{-2}p^T Bp\alpha$;
(154) $\nabla^2_{pp}c(p) = (\alpha^T p)^{-1}B - (\alpha^T p)^{-2}Bp\alpha^T - (\alpha^T p)^{-2}\alpha p^T B + (\alpha^T p)^{-3}\, p^T Bp\alpha\alpha^T$.

We now prove that the c(p) defined by (146)-(149) (with α predetermined) is a flexible functional form at the point p^*. Using the restrictions (149), $Bp^* = 0_N$, we have $p^{*T}Bp^* = p^{*T}0_N = 0$. Thus evaluating (153) and (154) at $p = p^*$ yields the following equations:

(155) $\nabla_p c(p^*) = b$;
(156) $\nabla^2_{pp}c(p^*) = (\alpha^T p^*)^{-1}B$.

We need to satisfy equations (91) and (92) above to show that the c(p) defined by (146)-(149) is flexible at p^*. Using (155), we can satisfy equations (91) if we choose b as follows:

(157) $b \equiv \nabla c^*(p^*)$.

Using (156), we can satisfy equations (92) by choosing B as follows:

(158) $B \equiv (\alpha^T p^*)^{-1}\nabla^2 c^*(p^*)$.

Since $\nabla^2 c^*(p^*)$ is a symmetric matrix, B will also be a symmetric matrix and so the symmetry restrictions (148) will be satisfied for the B defined by (158). Moreover, since $c^*(p)$ is assumed to be a linearly homogeneous function, Euler's Theorem implies that

(159) $\nabla^2 c^*(p^*)p^* = 0_N$.

Equations (158) and (159) imply that the B defined by (158) satisfies the linear restrictions (149). This completes the proof of the flexibility property for the normalized quadratic unit cost function.

It is convenient to define the vector of *normalized input prices*, $v^T \equiv [v_1,...,v_N]$ as follows:

(160) $v \equiv (p^T\alpha)^{-1}p$.

The system of input demand functions x(y,p) that corresponds to the normalized quadratic unit cost function c(p) defined by (146) can be obtained using Shephard's Lemma in the usual way:

(161) $x(y,p) = y\nabla c(p)$.

Using (161) and definition (146) evaluated at the period t data, we obtain the following system *of estimating equations*:

(162) $x^t/y^t = b + Bv^t - (1/2)v^{tT}Bv^t\alpha + e^t$; $\qquad t = 1,...,T$

where x^t is the observed period t input vector, y^t is the period t output, $v^t \equiv p^t/\alpha^T p^t$ is the vector of period t normalized input prices and $e^t \equiv [e_1^t,...,e_N^t]^T$ is a vector of stochastic error terms. Equations (162) can be used in order to statistically estimate the parameters in the b vector and the B matrix. Note that equations (162) are linear in the unknown parameters. Note also that the symmetry restrictions (148) can be imposed when estimating the system of equations (162) or their validity can be tested.

Once estimates for b and B have been obtained (denote these estimates by b^* and B^* respectively), then equations (162) can be used in order to generate a period t vector of fitted input demands, x^{t*} say:

(163) $x^{t*} \equiv y^t[b^* + B^*v^t - (1/2)v^{tT}B^*v^t\alpha]$; $\qquad t = 1,...,T$.

Equations (154) and (161) may be used in order to calculate the matrix of period t *estimated input price derivatives*, $\nabla_p x(y^t,p^t) = \nabla^2_{pp}C(y^t,p^t)$. The estimated matrix of second order partial derivatives $\nabla^2_{pp}C(y^t,p^t)$ for t = 1,...,T is the following one:

(164) $[C_{ij}^{t*}] \equiv y^t[(\alpha^T p^t)^{-1}B^* - (\alpha^T p^t)^{-2}B^*p^t\alpha^T - (\alpha^T p^t)^{-2}\alpha p^{tT}B^* + (\alpha^T p^t)^{-3}\, p^{tT}B^*p^t\alpha\alpha^T]$.

Equations (163) and (164) may be used in order to obtain estimates for the matrix *of period t input demand price elasticities*, $[E_{ij}^t]$:

(165) $E_{ij}^t \equiv \partial \ln x_i(y^t,p^t)/\partial \ln p_j = p_j^t\, C_{ij}^{t*}/x_i^{t*}$; $\qquad$ i,j = 1,...,N ; t = 1,...,T

where x_i^{t*} is the ith component of the vector of fitted demands x^{t*} defined by (163).

There is one important additional topic that we have to cover in our discussion of the normalized quadratic functional form: what conditions on b and B are necessary and sufficient to ensure that c(p) defined by (146)-(149) is concave in the components of the price vector p?

The function c(p) will be concave in p if and only if $\nabla^2 c(p)$ is a negative semidefinite matrix for each p in the domain of definition of c. Evaluating (154) at $p = p^*$ and using the restrictions (149) yields:

(166) $\nabla^2 c(p^*) = (\alpha^T p^*)^{-1} B$.

Since $\alpha > 0_N$ and $p^* >> 0_N$, $\alpha^T p^* > 0$. Thus in order for c(p) to be a concave function of p, the following necessary condition must be satisfied:

(167) B is a negative semidefinite matrix.

We now show that the *necessary condition* (167) is also *sufficient* to imply that c(p) is concave over the set of p such that $p >> 0_N$. Unfortunately, the proof is somewhat involved.[61]

Let $p >> 0_N$. We assume that B is negative semidefinite and we want to show that $\nabla^2 c(p)$ is negative semidefinite or equivalently, that $-\nabla^2 c(p)$ is positive semidefinite. Thus for any vector z, we want to show that $-z^T \nabla^2 c(p) z \geq 0$. Using (154), this inequality is equivalent to:

(168) $-(\alpha^T p)^{-1} z^T B z + (\alpha^T p)^{-2} z^T B p \alpha^T z + (\alpha^T p)^{-2} z^T \alpha p^T B z - (\alpha^T p)^{-3} p^T B p z^T \alpha \alpha^T z \geq 0$
or
(169) $-(\alpha^T p)^{-1} z^T B z - (\alpha^T p)^{-3} p^T B p (\alpha^T z)^2 \geq -2(\alpha^T p)^{-2} z^T B p \alpha^T z$ $\qquad$ using $B = B^T$.

Define $A \equiv -B$. Since B is symmetric and negative semidefinite by assumption, A is symmetric and positive semidefinite. Thus there exists an orthonormal matrix U such that

(170) $U^T A U = \Lambda$;
(171) $U^T U \;\; = I_N$

where I_N is the N by N identity matrix and Λ is a diagonal matrix with the nonnegative eigenvalues of A, λ_i, i = 1,...,N, running down the main diagonal. Now premultiply both

[61] The proof is due to Diewert and Wales (1987; 66).

sides of (170) by U and postmultiply both sides by U^T. Using (171), $U^T = U^{-1}$, and the transformed equation (170) becomes the following equation:

$$\begin{aligned}(172)\ A &= U\Lambda U^T \\ &= U\Lambda^{1/2}\,\Lambda^{1/2}\,U^T \\ &= U\Lambda^{1/2}\,U^TU\,\Lambda^{1/2}\,U^T \qquad \text{since } U^TU = I_N \\ &= SS\end{aligned}$$

where $\Lambda^{1/2}$ is the diagonal matrix that has the nonnegative square roots $\lambda_i^{1/2}$ of the eigenvalues of A running down the main diagonal and the symmetric square root of A matrix S is defined as

$$(173)\ S \equiv U\Lambda^{1/2}\,U^T.$$

If we replace – B in (169) with A, the inequality that we want to establish becomes

$$(174)\ 2(\alpha^Tp)^{-1}\,z^TAp\alpha^Tz \le z^TAz + (\alpha^Tp)^{-2}\,p^TAp(\alpha^Tz)^2$$

where we have also multiplied both sides of (169) by the positive number α^Tp in order to derive (174) from (169).

Recall the Cauchy-Schwarz inequality for two vectors, x and y:

$$(175)\ x^Ty \le (x^Tx)^{1/2}(y^Ty)^{1/2}.$$

Now we are ready to establish the inequality (174). Using (172), we have:

$$\begin{aligned}(176)\ (\alpha^Tp)^{-1}\,z^TAp\alpha^Tz &= (\alpha^Tp)^{-1}\,z^TSSp\alpha^Tz \\ &\le (z^TSS^Tz)^{1/2}\,([\alpha^Tp]^{-2}[\alpha^Tz]^2p^TS^TSp)^{1/2} \\ &\qquad \text{using (175) with } x^T \equiv z^TS \text{ and } y \equiv (\alpha^Tp)^{-1}(\alpha^Tz)\,Sp \\ &= (z^TSSz)^{1/2}([\alpha^Tp]^{-2}[\alpha^Tz]^2p^TSSp)^{1/2} \qquad \text{using } S = S^T \\ &= (z^TAz)^{1/2}([\alpha^Tp]^{-2}[\alpha^Tz]^2\,p^TAp)^{1/2} \qquad \text{using (172), } A = SS \\ &\le (1/2)(z^TAz) + (1/2)[\alpha^Tp]^{-2}[\alpha^Tz]^2(p^TAp)\end{aligned}$$

where the last inequality follows using the nonnegativity of z^TAz, p^TAp, the positivity of α^Tz and the Theorem of the Arithmetic and Geometric Mean.[62]

The inequality (176) is equivalent to the desired inequality (174).

Thus the normalized quadratic unit cost function defined by (146)-(149) will be concave over the set of positive prices if and only if the symmetric matrix B is negative semidefinite. Thus after econometric estimates of the elements of B have been obtained using the system of estimating equations (162), we need only check that the resulting estimated B^* matrix is negative semidefinite.

[62] This proof is due to Diewert and Wales (1987).

However, suppose that the estimated B^* matrix is *not* negative semidefinite. How can one reestimate the model, impose negative semidefiniteness on B, but without destroying the flexibility of the normalized quadratic functional form?

The desired imposition of negative semidefiniteness can be accomplished using a technique due to Wiley, Schmidt and Bramble (1973): simply replace the matrix B by

(177) $B \equiv - AA^T$

where A is an N by N lower triangular matrix; i.e., $a_{ij} = 0$ if $i < j$.[63]

We also need to take into account the restrictions (149), $Bp^* = 0_N$. These restrictions on B can be imposed if we impose the following restrictions on A:

(178) $A^Tp^* = 0_N$.

To show how this curvature imposition technique works, let $p^* = 1_N$ and consider the case N = 2. In this case, we have:

$A \equiv \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$ and $A^T = \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix}$.

The restrictions (178) become: $A^T 1_2 = \begin{bmatrix} a_{11} + a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

and hence we must have $a_{21} = - a_{11}$ and $a_{22} = 0$. Thus in this case,

(179) $B \equiv - AA^T = - \begin{bmatrix} a_{11} & 0 \\ -a_{11} & 0 \end{bmatrix} \begin{bmatrix} a_{11} & -a_{11} \\ 0 & 0 \end{bmatrix} = - \begin{bmatrix} a_{11}^2 & -a_{11}^2 \\ -a_{11}^2 & a_{11}^2 \end{bmatrix} = a_{11}^2 \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$.

Equations (179) show how the elements of the B matrix can be defined in terms of the single parameter, a_{11}^2. Note that with this reparameterization of the B matrix, it will be necessary to use nonlinear regression techniques rather than modifications of linear regression techniques. This turns out to be the cost of imposing the correct curvature conditions on the unit cost function.

11. The Konüs Byushgens Fisher Unit Cost Function

---

[63] Since $z^TAA^Tz = (A^Tz)^T(A^Tz) = y^Ty \geq 0$ for all vectors z, AA^T is positive semidefinite and hence $- AA^T$ is negative semidefinite. Diewert and Wales (1987; 53) showed that any positive semidefinite matrix can be written as AA^T where A is lower triangular. Hence, it is not restrictive to reparameterize an arbitrary negative semidefinite matrix B as $- AA^T$.

Define the KBF unit cost function, c(p), as follows:[64]

(180) $c(p) \equiv (p^T Bp)^{1/2}$; $B = B^T$

where B is an N by N symmetric matrix which has one positive eigenvalue (with a strictly positive eigenvector) and the remaining N−1 eigenvalues are negative or zero. The vector of first order partial derivatives of this unit cost function, $\nabla c(p)$, and the matrix of second order partials, $\nabla^2 c(p)$, are equal to the following expressions:

(181) $\nabla c(p) = Bp/(p^T Bp)^{1/2}$;
(182) $\nabla^2 c(p) = (p^T Bp)^{-1/2}\{B - Bp(p^T Bp)^{-1} p^T B\}$.

At this point, we need to determine the region of price space where the c(p) defined by (180) is a concave function. In general, the unit cost function defined by (180) will not be concave for all strictly positive price vectors p.[65] In order for a unit cost function to provide a valid global representation of homothetic preferences, it must be a nondecreasing, linearly homogeneous and concave function over the positive orthant. However, in order for c to provide a valid local representation of preferences, we need only require that c(p) be positive, nondecreasing, linearly homogeneous and concave over a convex subset of prices, say S, where S has a nonempty interior.[66] It is obvious that c(p) defined by (20) is linearly homogeneous. The nondecreasing property will hold over S if the gradient vector $\nabla c(p)$ defined by (181) is strictly positive for $p \in S$ and the concavity property will hold if $\nabla^2 c(p)$ defined by (182) is a negative semidefinite matrix for $p \in S$. We will show how the regularity region S can be determined shortly but first, we will indicate why the c(p) defined by (20) is a flexible functional form[67] since this explanation will help us to define an appropriate region of regularity.

Let $p^* \gg 0_N$ be a strictly positive reference price vector and suppose that we are given an arbitrary unit cost function $c^*(p)$ that is twice continuously differentiable in a neighbourhood around p^*.[68] Let $x^* \equiv \nabla c^*(p) \gg 0_N$ be the strictly positive vector of first order partial derivatives of $c^*(p^*)$ and let $S^* \equiv \nabla^2 c^*(p^*)$ be the negative semidefinite symmetric matrix of second order partial derivatives of c^* evaluated at p^*. Euler's

[64] This is a special case of a functional form due to Denny (1974), which Diewert (1976; 131) called the quadratic mean of order r unit cost function. This functional form with r = 2 was introduced into the economics literature by Konüs and Byushgens (1926; 168) and its connection to the Fisher (1922) ideal price index was explained by these authors and Diewert (1976). See Problem 22 below.

[65] The following analysis of the regularity conditions for the c(p) defined by (180) is due to Diewert and Hill (2010).

[66] See Blackorby and Diewert (1979) for more details on local representations of preferences using duality theory.

[67] Diewert (1976; 130) established the flexibility of c(p) defined by (180) as part of a more general result.

[68] Of course, in addition, we assume that c^* satisfies the appropriate regularity conditions for a unit cost function. Using Euler's Theorem on homogeneous functions, the fact that c^* is linearly homogeneous and differentiable at p^* means that the derivatives of c^* satisfy the following restrictions: $c^*(p^*) = p^{*T}\nabla c^*(p^*)$ and $\nabla^2 c^*(p^*)p^* = 0_N$. The unit cost function c defined by (180) satisfies analogous restrictions at $p = p^*$. These restrictions simplify the proof of the flexibility of c at the point p^*.

Theorem on homogeneous functions implies that S^* satisfies the following matrix equation:

(183) $S^*p^* = 0_N$.

In order to establish the flexibility of the KBF c defined by (180), we need only show that there are enough free parameters in the B matrix so that the following equations are satisfied:

(184) $\nabla c(p^*) = x^*$;
(185) $\nabla^2 c(p^*) = S^*$.

In order to prove the flexibility of c, it is convenient to reparameterize the B matrix. Thus we now set B equal to:

(186) $B = bb^T + A$

where $b >> 0_N$ is a positive vector and A is a negative semidefinite matrix which has rank equal to at most N−1 and it satisfies the following restrictions:

(187) $Ap^* = 0_N$.

Note that bb^T is a rank one positive semidefinite matrix with $p^{*T}bb^Tp^* = (b^Tp^*)^2 > 0$ and A is a negative semidefinite matrix and satisfies $p^{*T}Ap^* = 0$. Thus it can be seen that B is a matrix with one positive eigenvalue and the other eigenvalues are negative or zero.

Substitute (181) into (184) in order to obtain the following equation:

$$\begin{aligned}(188)\ x^* &= Bp^*/(p^{*T}Bp^*)^{1/2} \\ &= [bb^T + A]p^*/(p^{*T}[bb^T + A]p^*)^{1/2} && \text{using (186)} \\ &= bb^Tp^*/(p^{*T}bb^Tp^*)^{1/2} && \text{using (187)} \\ &= b.\end{aligned}$$

Thus if we choose b equal to x^*, equation (184) will be satisfied. Now substitute (182) into (183) and obtain the following equation:

$$\begin{aligned}(189)\ S^* &= (p^{*T}Bp^*)^{-1/2}\{B - Bp^*(p^{*T}Bp^*)^{-1}p^{*T}B\} \\ &= (p^{*T}bb^Tp^*)^{-1/2}\{bb^T + A - bb^Tp^*(p^{*T}bb^Tp^*)^{-1}p^{*T}bb^T\} && \text{using (186) and (187)} \\ &= (b^Tp^*)^{-1}A && \text{using } b^Tp^* > 0.\end{aligned}$$

Thus if we choose A equal to $(b^Tp^*)S^*$, equation (185) will be satisfied and the flexibility of c defined by (180) is established.[69]

[69] We need to check that A is negative semidefinite (which it is since it is a positive multiple of the negative semidefinite substitution matrix S^*) and that A satisfies the restrictions in (187), since we used these restrictions to derive (188) and the second line in (189). But A does satisfy (187) since A satisfies (183).

Now we are ready to define the region of regularity for c defined by (180).[70] Consider the following set of prices:

(190) $S \equiv \{p : p >> 0_N ; Bp >> 0_N\}$.

If $p \in S$, then it can be seen that $c(p) = (p^TBp)^{1/2} > 0$ and using (181), $\nabla c(p) >> 0_N$. However, it is more difficult to establish the concavity of c(p) over the set S. We first consider the case where the matrix B has full rank so that it has one positive eigenvalue and N−1 negative eigenvalues. Let $p \in S$ and using equation (182), we see that $\nabla^2 c(p)$ will be negative semidefinite if and only if the matrix M defined as:

(191) $M \equiv B - Bp(p^TBp)^{-1}p^TB$

is negative semidefinite. Note that M is equal to the matrix B plus the rank 1 negative semidefinite matrix $- Bp(p^TBp)^{-1}p^TB$. B has one positive eigenvalue and the remaining eigenvalues are 0 or negative. Since M is B plus a negative semidefinite matrix, the eigenvalues of M cannot be greater than the eigenvalues of B. Now consider two cases; the first case where B has one positive and N−1 negative eigenvalues and the second case where B has N−1 negative or zero eigenvalues in addition to its positive eigenvalue. Consider case 1, let $p \in S$ and calculate Mp:

(192) $Mp = [B - Bp(p^TBp)^{-1}p^TB]p = 0_N$.

The above equation shows that $p \neq 0_N$ is an eigenvector of M that corresponds to a 0 eigenvalue. Now the addition of a negative semidefinite matrix to B can only make the N−1 negative eigenvalues of B more negative (or leave them unchanged) so we conclude that the addition of the negative semidefinite matrix $- Bp(p^TBp)^{-1}p^TB$ to B has converted the positive eigenvalue of B into a zero eigenvalue and hence M is negative semidefinite.

Case 2 follows using a perturbation argument.

Thus we have shown that the KBF unit cost function c(p) defined by (180) is positive, increasing in the components of p and concave in p over the region of prices S defined by (190).

It is useful to show if $c(p) \equiv (p^TBp)^{1/2}$ is defined by (180), then we can decompose the matrix B into $bb^T + A$ where $b >> 0_N$ and A is a negative semidefinite matrix with $Ap^* = 0_N$ for some $p^* >> 0_N$. Recall that definition (180) specified that $c(p) \equiv (p^TBp)^{1/2}$ where B is an N by N symmetric matrix which has one positive eigenvalue (with a strictly positive eigenvector) and the remaining N−1 eigenvalues are negative or zero. Let $\lambda_1 > 0$ and $\lambda_i \leq 0$ for i = 2,3,...,N be the eigenvalues of B and let the column vectors u^i be the corresponding eigenvectors, which are orthonormal to each other; i.e., $u^{iT}u^i = 1$ for i = 1,...,N and $u^{iT}u^j = 0$ for all $i \neq j$. Then it is well known that the matrix B has the following representation:

[70] The region of regularity can be extended to the closure of the set S.

(193) $B = \Sigma_{i=1}^{N} \lambda_i u^i u^{iT}$.

Using the regularity conditions in definition (180), it can be seen that the first eigenvector u^1 is strictly positive. Make the following definitions:

(194) $p^* \equiv u^1 >> 0_N$; $b \equiv (\lambda_1)^{1/2} u^1$; $A \equiv \Sigma_{i=2}^{N} \lambda_i u^i u^{iT}$.

It can be seen that A is a negative semidefinite matrix. Since $u^1 = p^*$ is orthogonal to $u^2,...,u^N$, $Ap^* = 0_N$. Thus we have $B = bb^T + A$ where b is a positive vector and A is negative semidefinite with $Ap^* = 0_N$.

The following problems show the connection of the KBF functional form with Irving Fisher's (1922) ideal index number formula.

Problems

21. Suppose that a producer's unit cost function is defined by (180). Assume cost minimizing behavior on the part of the producer for periods 1 and 2 so that using Shephard's Lemma, we have:

(i) $x^t = \nabla c(p^t) y^t$; $\quad$ t = 1,2

where p^t, x^t and y^t are the period t input price and quantity vectors and y^t is the period t output level for t = 1,2. (a) Show that

(ii) $x^t / p^{tT} x^t = \nabla c(p^t) / c(p^{it})$; $\quad$ t = 1,2.

(b) Show that we also have the following equations:

(iii) $x^t / p^{tT} x^t = Bp^t / c(p^t)$; $\quad$ t = 1,2.

22. Continuation of 21: The *Fisher* (1922) *ideal input price index* P_F is defined as the following function of the observed input price and quantity vectors for periods 1 and 2:

(i) $P_F(p^1,p^2,x^1,x^2) \equiv [p^{2T}x^1 \ p^{2T}x^2 / p^{1T}x^1 \ p^{1T}x^2]^{1/2}$.

Assume that p^1,p^2,x^1,x^2 satisfy equations (i) in Problem 21 where the KBF unit cost function c(p) is defined by (180). Show that

(ii) $P_F(p^1,p^2,x^1,x^2) = c(p^2)/c(p^1)$.

Hint: Note that the inner products of p^2 with $x^1/p^{1T}x^1$ and p^1 with $x^2/p^{2T}x^2$ appear in the formula (i) above for $P_F(p^1,p^2,x^1,x^2)$. Apply part (b) of Problem 21.

Comment: The ratio of unit costs, $c(p^2)/c(p^1)$, can be interpreted as a theoretical input price index, (due originally to Konüs (1924) in the consumer context). Equation (ii) above tells us that this theoretical input cost index can be calculated using just observed input price and quantity data for the two periods under consideration using the Fisher index provided that the producer is cost minimizing in the two periods and has the production function that is dual to the unit cost function defined by (180). Thus no econometric estimation is necessary in order to construct the ratio of unit costs.[71]

We conclude this section by looking at the problems associated with estimating the unknown parameters in the symmetric B matrix, assuming that we have data on a production unit producing one output and using N inputs for T time periods. Using (181), Shephard's Lemma and definition (180) evaluated at the period t data, we obtain the following system *of estimating equations*:

$$(195)\ x^t/y^t = Bp^t/(p^{tT}Bp^t)^{1/2} + e^t\ ; \qquad t = 1,\ldots,T$$

where x^t is the observed period t input vector, y^t is the period t output, p^t is the vector of period t input prices and $e^t \equiv [e_1^t,\ldots,e_N^t]^T$ is a vector of stochastic error terms with 0 means. Equations (195) can be used in order to statistically estimate the N(N+1)/2 independent b_{ij} parameters in the B matrix. However, the system of equations defined by (195) is nonlinear in the unknown parameters. Define period t unit cost by $c^t \equiv p^{tT}x^t/y^t$. In theory, c^t should equal $(p^{tT}Bp^t)^{1/2}$ plus an error term. Thus the system of estimating equations (195) can be replaced by the following system:

$$(196)\ c^tx^t/y^t = Bp^t + e^{t*}\ ; \qquad t = 1,\ldots,T$$

where $e^{t*} \equiv [e_1^{t*},\ldots,e_N^{t*}]^T$ is a new vector of stochastic error terms with 0 means. Note that the new system of estimating equations defined by (196) is linear in the unknown b_{ij}.[72]

As was the case when estimating the normalized quadratic unit cost function, it will often turn out that the estimated B matrix will not satisfy the regularity conditions that are associated with definition (180). As we have seen above, B may be estimated as the equivalent expression equal to $bb^T + A$ where b is a strictly positive vector and A is a symmetric negative semidefinite matrix with $Ap^* = 0_N$ for some strictly positive reference vector p^*. Thus we need only set $A = -\,CC^T$ where C is a lower triangular matrix with

---

[71] This result is much more important in the consumer context where we interpret f(x) as a utility function defined over consumption vectors x and c(p) is the dual unit expenditure function. Note that utility cannot be observed whereas output can be observed.

[72] In the consumer context where output y^t is replaced by (unobservable) utility u^t and x^t is the period t consumption vector, rewrite (195) as $x^t = u^tBp^t/(p^{tT}Bp^t)^{1/2}$ where we have dropped the error terms. Total period t expenditure is $p^{tT}x^t = u^t(p^{tT}Bp^t)^{1/2}$. Thus we obtain $x^t/p^{tT}x^t = Bp^t/p^{tT}Bp^t$. Premultiply both sides of equation n by p_n^t and we obtain the following system of estimating equations: $p_n^tx_n^t/p^{tT}x^t \equiv s_n^t = p_n^t\Sigma_{i=1}^N b_{ni}p_i^t/p^{tT}Bp^t + e_n^t$ for n = 1,...,N and t = 1,...,T. We need to impose a normalization on the elements of the B matrix such as $b_{11} = 1$ and since we have share equations, we need to drop one of these share equations in the nonlinear estimation procedure. For an example of this methodology in the consumer context, see Diewert and Feenstra (2017).

$C^T p^* = 0_N$ and the correct curvature conditions will be imposed on the resulting functional form for the unit cost function defined as follows:

$$(197)\ c(p) \equiv (p^T[bb^T - CC^T]p)^{1/2} = (p^T[bb^T - \Sigma_{i=1}^{N-1} c^i c^{iT}]p)^{1/2}$$

where $c^{1T} \equiv [c_1{}^1,c_2{}^1,...,c_N{}^1]$, $c^{2T} \equiv [0,c_2{}^2,...,c_N{}^2]$, $c^{3T} \equiv [0,0,c_3{}^3,...,c_N{}^3]$, ..., $c^{(N-1)T} \equiv [0,...,0,c_{N-1}{}^{N-1},c_N{}^{N-1}]$ and $c^{nT}p^* = 0$ for n = 1,2,...,N−1.

We have considered four flexible functional forms for a unit cost function: the Generalized Leontief, the translog, the normalized quadratic and the KBF functional forms. The last two functional forms have the advantage that concavity can be imposed on these functional forms without destroying the flexibility of the resulting functions. The normalized quadratic functional form has the disadvantage that it is usually necessary to choose the vector α [73] whereas all of the parameters for the KBF functional form can be estimated endogenously.

12. Semiflexible Functional Forms

In models where the number of commodities N is large, it can be difficult to estimate all of the parameters for a flexible functional form. Thus when estimating the parameters for the normalized quadratic defined by (146) above, it was necessary to estimate the elements of the N by N symmetric matrix B and for the KBF functional form, it was necessary to estimate the elements of the N by N symmetric matrix A in (186). If we impose concavity on these functional forms, then in both of these cases, the B and A matrices are replaced by $-CC^T$ where C is lower triangular and $Cp^* = 0_N$ for a reference positive price vector p^*. An effective way to estimate the C matrix is to estimate it one column at a time. Thus consider our estimating equations (162) for the Normalized Quadratic unit cost function. Replace the B matrix in these equations by $-CC^T$ where C is lower triangular and $Cp^* = 0_N$ and we obtain the following system of equations:

$$(198)\ x^t/y^t = b - CC^Tv^t + (1/2)v^{tT}CC^Tv^t\alpha + e^t\ ; \qquad t = 1,...,T.$$

In Stage 1, we set $C = 0_{N\times N}$ and use the resulting equations in (198) in order to estimate the vector of parameters b. In Stage 2, set $CC^T = c^1c^{1T}$ where $c^{1T} \equiv [c_1{}^1,c_2{}^1,...,c_N{}^1]$ and $c^{1T}p^* = 0$. Equations (198) now become a nonlinear regression model. For starting parameter values, use the b vector that was estimated in Stage 1 and set the vector $c^1 = 0_N$. In Stage 3, set $CC^T = c^1c^{1T} + c^2c^{2T}$ where $c^{1T} \equiv [c_1{}^1,c_2{}^1,...,c_N{}^1]$, $c^{2T} \equiv [0,c_2{}^2,...,c_N{}^2]$ and $c^{iT}p^* = 0$ for i = 1,2. For starting parameter values, use the b and c^1 vectors that were estimated in Stage 2 and set the vector $c^2 = 0_N$. This procedure of gradually adding nonzero columns of the lower triangular C matrix can be continued until the full number of N−1 nonzero columns have been added, provided that the number of time series observations T is large enough compared to N, the number of commodities in the

[73] Thus different choices for the α vector could lead to different estimates for elasticities of demand. N−1 components of the α vector could be estimated along with the remaining parameters but then we would not have a parsimonious flexible functional form.

model.[74] However, in models where T is small relative to N, the above procedure of adding nonzero columns to A will have to be stopped well before the maximum number of N−1 nonzero columns has been added, due to the lack of degrees of freedom. Suppose that we stop the above procedure after $K < N-1$ nonzero columns have been added. Then Diewert and Wales (1988; 330) called the resulting normalized quadratic functional form a *flexible of degree K* functional form or a *semiflexible functional form*. A flexible of degree K functional form for a cost function can approximate an arbitrary twice continuously differentiable functional form to the second order at some point, except the matrix of second order partial derivatives of the functional form with respect to prices is restricted to have maximum rank K instead of the maximum possible rank, N−1.

The cost of using a semiflexible functional form of degree K where K is less than N−1 is that we will miss out on the part of CC^T that corresponds to the smallest eigenvalues of this matrix. In many situations, this cost will be very small; i.e., as we go through the various stages of estimating C by adding an extra nonzero column to C at each stage, we can monitor the increase in the final log likelihood (if we use maximum likelihood estimation) and when the increase in Stage k+1 over Stage k is "small", we can stop adding extra columns, secure in the knowledge that we are not underestimating the size of CC^T by a large amount.

This semiflexible technique has not been widely applied but it would seem to offer some big advantages in estimating substitution matrices in situations where there are a large number of commodities in the model.[75]

13. The Use of Splines for Modeling Technical Progress.

Recall the definitions for the Generalized Leontief, normalized quadratic and KBF unit cost functions c(p) given by (97), (146) and (180). If these functions are estimated in the time series context for a production unit for say T periods, then a problem will often occur: these functional forms make no allowance for *technical progress* that may have taken place over the sample time period. This problem can be solved if we add the function d^Tpt to the unit cost function c(p) where $d^T \equiv [d_1,...,d_N]$ is an N dimensional vector of technical progress parameters and t is a scalar time variable which takes on the value t for time period t. Thus choose a flexible functional form for the unit cost function c(p) and add the function d^Tpt to it. Using our usual notation for a data set on inputs x^t, input prices p^t and output levels y^t for period t, we obtain the following system of estimating equations using Shephard's Lemma:

(199) $x^t/y^t = \nabla c(p^t) + dt + e^t$; $\qquad t = 1,...,T$

[74] In empirical applications, typically a final stage $K < N-1$ will be reached where the addition of another column to the CC^T matrix leads to no increase in log likelihood and the last column c^K is a column of zeros.

[75] Diewert and Lawrence in some unpublished work have successfully estimated semiflexible models for profit functions for 40 commodities. Neary (2004) used semiflexible functional forms for 11 commodity groups.

where e^t is a suitable error vector. If we choose c(p) to be the normalized quadratic unit cost function, then the resulting estimating equations (199) will be linear in the unknown parameters.[76]

However, in many applications of this model, the results may not be satisfactory. The problem with the model defined by equations (199) is that the resulting measures of technical progress are too smooth; i.e., typically if one looks at the residuals generated by the model, substantial amounts of autocorrelation will be present in the estimating equations. This is an indication that rates of technical progress are not constant over the sample time period. Under these circumstances, it will be useful to replace the function simple linear function d^Tpt by the following piece-wise linear spline function, $\tau(p,t)$, defined as follows:

$$\begin{aligned}(200)\ \tau(p,t) &\equiv d^{1T}pt \text{ if } 1 \leq t \leq t^{1*} ;\\ &\equiv d^{1T}pt^{1*} + d^{2T}p(t-t^{1*}) \text{ if } t^{1*} \leq t \leq t^{2*} ;\\ &\equiv d^{1T}pt^{1*} + d^{2T}p(t^{2*}-t^{1*}) + d^{3T}p(t-t^{2*}) \text{ if } t^{2*} \leq t \leq T\end{aligned}$$

where d^1, d^2 and d^3 are N dimensional technical progress parameters and t^{1*} and $t^{2*} > t^{1*}$ are two time periods where the piece-wise linear function of time t, $\tau(p,t)$, changes from one set of rates of technical progress to another set.[77] The estimating equations are now the following ones:[78]

$$\begin{aligned}(201)\ x^t/y^t &= \nabla c(p^t) + d^1t + e^t ; && 1 \leq t \leq t^{1*};\\ &= \nabla c(p^t) + d^1t^{1*} + d^2(t-t^{1*}) + e^t ; && t^{1*} < t \leq t^{2*};\\ &= \nabla c(p^t) + d^1t^{1*} + d^2(t^{2*}-t^{1*}) + d^3(t-t^{2*}) + e^t ; && t^{2*} < t \leq T.\end{aligned}$$

If we chose c(p) to be the normalized quadratic unit cost function, then, assuming that it is not necessary to impose concavity, the above estimating equations will be linear in the unknown parameters. For an example of the use of the above spline methodology, see Fox (1998).[79]

The above spline methodology for modeling technical progress can be modified to model nonconstant returns to scale technologies; see Fox and Grafton (2000).

The above linear spline model has the disadvantage that *rates* of technical progress will typically jump in a discontinuous manner as we move from one linear spline segment to the following one. This problem can be remedied (at the cost of a more complicated set

---

[76] However, if we impose concavity on the normalized quadratic functional form, then the resulting estimating equations will be nonlinear in the unknown parameters. For a worked example of this methodology for modelling technical progress, see Diewert and Wales (1987).
[77] The break points t^{1*} and t^{2*} can be chosen by running a preliminary regression of the form (199) and examining the regression residuals to see when these turning points occur. In our example, we have three time periods where the rates of technical progress are linear in time. If necessary, additional break points can be added at the cost of having to estimate additional parameter vectors d^i.
[78] If the unit cost function is translog, then the estimating equations will be somewhat different.
[79] Fox used a more scientific method to pick the break points (cross validation).

of estimating equations) if the *linear* splines in time t are replaced with *quadratic* splines in t. For an example of the quadratic spline approach, see Diewert and Wales (1992).

14. Allowing for Flexibility at Two Sample Points

There can be a problem with our two flexible functional forms for unit cost functions where the correct curvature conditions can be imposed (the normalized quadratic and the KBF unit cost functions): the elasticities of input demand that these functions generate in the time series context can exhibit substantial trends.

We need to derive a formula for the elasticity of demand for input n with respect to a change in the price of input k, say $E_{nk}(y,p)$ where y is output and p is an input price vector. Recall that the normalized quadratic unit cost function was defined by $c(p) \equiv b^Tp + (1/2)p^TBp/\alpha^Tp$ where α is predetermined and B is a symmetric matrix which satisfies $Bp^* = 0_N$.[80] The vector of first order partial derivatives and the matrix of second order partial derivatives of this unit cost function are as follows:

(202) $\nabla c(p) = b + Bv - (1/2)\, v^TBv\alpha$;
(203) $\nabla^2 c(p) = (\alpha^Tp)^{-1}[\, B - Bv\alpha^T - \alpha v^TB + v^TBv\alpha\alpha^T]$

where $v \equiv p/\alpha^Tp$ is a vector of normalized input prices. The system of input demand functions that is generated by this functional form is $x(y,p) \equiv y\nabla c(p)$ and the N by N matrix of input demand derivatives with respect to input prices is $\nabla_p x(y,p) \equiv y\nabla^2 c(p)$. Using (203), we see that the elasticity $E_{nk}(y,p) \equiv [p_k/x_n]\partial x_n(y,p)/\partial p_k$ is equal to the following expression:

(204) $E_{nk}(y,p) = [p_k/\alpha^Tp][y/x_n][b_{nk} - B_{n\bullet}v\alpha_k - B_{k\bullet}v\alpha_n + v^TBv\alpha_n\alpha_k]$; n, k = 1,...,N

where b_{nk} is the nk^{th} element of the matrix B, $B_{i\bullet}$ denotes the ith row of the B matrix for i = 1,...,N and $v \equiv p/\alpha^Tp$ is the vector of normalized prices; i.e., the components of the input price vector p are divided by α^Tp. Note that when $p = p^*$, the restrictions imply that $v^{*T}Bv^* = 0$ and $B_{i\bullet}v^* = 0$ for i = 1,...,N where $v^* \equiv p^*/\alpha^Tp^*$. Thus

(205) $E_{nk}(y,p^*) = [p_k^*/\alpha^Tp^*][y/x_n]b_{nk}$; n, k = 1,...,N.

The reference price vector p^* will usually be a representative input price vector for the sample under consideration. Thus the price elasticity of input demand when evaluated at these reference prices, $E_{nk}(y,p^*)$, will be equal to the constant term b_{nk} times the price ratio term p_k^*/α^Tp^* times the quantity ratio term y/x_n. The remaining 3 terms on the right hand side of (204) will be equal to zero when $p = p^*$. Thus the first term will generally be the most significant term that defines $E_{nk}(y,p)$ for a general input price vector. If there are substantial divergent trends in either input prices p or input quantities x, it can be seen

[80] See equations (146)-(149) above. We also require that B be negative semidefinite, a property which can be imposed as was explained in Section 10 above.

that $[p_k/\alpha^Tp][y/x_n]b_{nk}$ will also have substantial trends and hence $E_{nk}(y^t,p^t)$ will, in general, also exhibit substantial trends under these conditions.

What can be done to remedy this problem of trending elasticities? If the number of observations $\tau+1$ is relatively large compared to the number of inputs N, then we can set the unit cost function equal to the following function of time t:

$$(206)\ c(p,t) \equiv (1-\tau^{-1}t)b^{1T}p + \tau^{-1}tb^{2T}p + (1/2)p^T[(1-\tau^{-1}t)B^1 + \tau^{-1}tB^2]p/\alpha^Tp\ ; \quad t = 0,1,2,...,\tau$$

where $B^1p^0 = 0_N$ and $B^2p^T = 0_N$.[81] Thus the resulting unit cost function evaluated at period 0 is $c(p,0) \equiv b^{1T}p + (1/2)p^TB^1p/\alpha^Tp$ and evaluated at period T is $c(p,T) \equiv b^{2T}p + (1/2)p^TB^2p/\alpha^Tp$; i.e., the resulting unit cost function is flexible at two data points. If there are trends in input demand elasticities using this functional form, then these trends are implied by the data rather than by the choice of the functional form.[82] Note that the unit cost function defined by (206) allows for biased technical change over the sample period; i.e., it allows for trends in the $b \equiv (1-t)b^1 + tb^2$ vector.[83]

It is possible to generalize the KBF unit cost function in a similar manner. Recall that this unit cost function was defined by (180): $c(p) \equiv (p^TBp)^{1/2}$ where $B \equiv bb^T + A$ and A is a negative semidefinite symmetric matrix which satisfies $Ap^* = 0_N$. The vector of first order partial derivatives was defined by (181). Using this equation and Shephard's Lemma, we have $x(y,p) = y\nabla c(p)$ and so $x/y = \nabla c(p)$. Thus using (181), we obtain the following equations:

$$(207)\ x(y,p) = y\nabla c(p) = yBp/(p^TBp)^{1/2}.$$

When $p = p^*$, using $B \equiv bb^T + A$ and $Ap^* = 0_N$, it can be seen that

$$(208)\ x^* \equiv x(y,p^*) = yb.$$

The matrix of input demand derivatives with respect to input prices is $\nabla_p x(y,p) = y\nabla^2 c(p)$. The matrix of second order partial derivatives of the unit cost function was defined by (182). Thus we have:

$$\begin{aligned}(209)\ \nabla_p x(y,p) &= y\nabla^2 c(p) \\ &= y(p^TBp)^{-1/2}\{B - Bp(p^TBp)^{-1}p^TB\} && \text{using (182)} \\ &= yc(p)^{-1/2}\{B - Bp(p^TBp)^{-1}p^TB\} && \text{since } c(p) \equiv (p^TBp)^{1/2} \\ &= yc(p)^{-1/2}\{B - y^{-2}x(y,p)x(y,p)^T\} && \text{using (207)}\end{aligned}$$

[81] We require that B^1 and B^2 be symmetric negative semidefinite matrices. If the estimated matrices fail to be negative semidefinite, then we can impose negative semidefiniteness by setting $B^i = -C^iC^{iT}$ for i = 1,2 where each C^i is an arbitrary lower triangular matrix satisfying $C^{1T}p^0 = 0_N$ and $C^{2T}p^T = 0_N$.
[82] This technique of imposing price flexibility at two points is due to Diewert and Lawrence (2002).
[83] If the residuals in the final model exhibit substantial autocorrelation, then it is possible to replace the b vector by a piece-wise linear function of time as was done in the previous section. This will allow for a more general pattern of technical change.

$$= yc(p)^{-1/2}\{bb^T + A - y^{-2}x(y,p)x(y,p)^T\} \qquad \text{using } B = bb^T + A.$$

Now evaluate (209) when $p = p^*$. We find that:

$$(210)\ \nabla_p x(y,p^*) = yc(p^*)^{-1/2}\{A + bb^T - y^{-2}x(y,p^*)x(y,p^*)^T\}$$
$$= yc(p^*)^{-1/2}\{A + bb^T - bb^T\} \qquad \text{using (208)}$$
$$= yc(p^*)^{-1/2}A.$$

Using (209), we see that the elasticity $E_{nk}(y,p) \equiv [p_k/x_n]\partial x_n(y,p)/\partial p_k$ is equal to the following expression:

$$(211)\ E_{nk}(y,p) = [p_k/c(p)][y/x_n(y,p)][a_{nk} + b_nb_k - y^{-2}x_n(y,p)x_k(y,p)]\ ; \qquad n, k = 1,...,N$$

where a_{nk} is the nkth element of the negative semidefinite matrix A (which satisfies $Ap^* = 0_N$), b_n is the nth element of the vector b and $x_n(y,p)$ is the nth element of the cost minimizing input vector x(y,p) defined by (207). Using (210), we see that when $p = p^*$, $E_{nk}(y,p^*) = [p_k/c(p^*)][y/x_n(y,p^*)]a_{nk}$ so that the last two terms on the right hand side of (211) sum to zero when $p = p^*$. Thus the first term associated with a_{nk} will generally be the most significant term that defines $E_{nk}(y,p)$ for a general input price vector. If there are substantial divergent trends in either input prices p or input quantities x, it can be seen that $[p_k/c(p^*)][y/x_n(y,p^*)]a_{nk}$ will also have substantial trends and hence $E_{nk}(y^t,p^t)$ will, in general, also exhibit substantial trends under these conditions.

Again, if the number of observations $\tau+1$ is relatively large compared to the number of inputs N, then we can set the KBF unit cost function equal to the following function of time t:

$$(212)\ c(p,t) \equiv (p^T[(1-\tau^{-1}t)b^{1T}p + \tau^{-1}tb^{2T}p + (1-\tau^{-1}t)A^1 + \tau^{-1}tA^2]p)^{1/2}\ ; \qquad t = 0,1,2,..., \tau$$

where $A^1p^0 = 0_N$ and $A^2p^T = 0_N$.[84] Thus the resulting unit cost function evaluated at period 0 is $c(p,0) \equiv (p^T[b^1b^{1T} + A^1]p)^{1/2}$ and evaluated at period T is $c(p,T) \equiv (p^T[b^1b^{1T} + A^1]p)^{1/2}$; i.e., the resulting unit cost function is flexible at two data points. As was the case for the normalized quadratic, if there are trends in input demand elasticities using this functional form, then these trends are implied by the data rather than by the choice of the functional form. Again, the unit cost function defined by (212) allows for biased technical change over the sample period; i.e., it allows for trends in the $b \equiv (1-t)b^1 + tb^2$ vector.

We turn our attention to multiple input and multiple output technologies.

15. National Product or Variable Profit Functions

[84] We require that A^1 and A^2 be symmetric negative semidefinite matrices. Again, if the estimated matrices fail to be negative semidefinite, then we can impose negative semidefiniteness by setting $A^i = -C^iC^{iT}$ for i = 1,2 where each C^i is an arbitrary lower triangular matrix satisfying $C^{1T}p^0 = 0_N$ and $C^{2T}p^T = 0_N$.

Up to now, we have only considered technologies that produce one output. In reality, production units (firms or industries) usually produce many outputs.[85] Hence, in this section, we consider technologies that produce many outputs while using many inputs.

Let S denote the technology set of a production unit. We decompose the inputs and outputs of the firm into two sets of commodities: *variable* and *fixed*. Let $y \equiv [y_1,\ldots,y_M]$ denote a vector of *variable net outputs* (if $y_m > 0$, then commodity m is an output while if $y_m < 0$, then commodity m is an input) and let $x \equiv [x_1,\ldots,x_N]$ denote a nonnegative vector of "fixed" inputs[86]. Thus the technology set S is a set of feasible variable net output and fixed input vectors, (y,x).

Let $p >> 0_M$ be a strictly positive vector of variable net output prices that the firm faces during a production period. Then conditional on a given vector of fixed inputs $x \geq 0_N$, we assume that the firm attempts to solve the following *conditional* or *variable profit maximization problem*:

(213) $\max_y \{p^Ty: (y,x)\in S\} \equiv \pi(p,x)$.

The optimized objective function, $\pi(p,x)$, has been called many names[87], depending on the context. Alternative names for this function are the *national product function* Samuelson (1953; 10), the *gross profit function* Gorman (1968), the *conditional profit function* McFadden (1966) (1978), the *variable profit function* Diewert (1973), the *GDP function* Kohli (1978) (1991) and the *value added function* Diewert (1978). If there are no intermediate inputs or imports in the outputs, then $\pi(p,x)$ becomes the *revenue function* Diewert (1974b). Some regularity conditions on the technology set S are required in order to ensure that the maximum in (213) exists. A simple set of sufficient conditions are:[88] (i) S is a closed set in R^{M+N} and (ii) for each $x \geq 0_N$, there exists a y such that $(y,x)\in S$ and the set of such y vectors is bounded from above. We will call these conditions the *minimal regularity conditions* on S.

Note that $\pi(p,x)$ is equal to the optimized objective function in (213) and is regarded as a function of the net output prices for variable commodities that the firm faces, p, as well as a function of the vector of fixed inputs, x, that the firm has at its disposal. Just as in section 2 above where we showed that the cost function C satisfied a number of regularity conditions without assuming much about the production function, we can now

[85] See Bernard, Redding and Schott (2010). In the sample of US firms considered by Hottman, Redding and Weinstein (2016; 1301), the mean number of products (measured by distinct barcodes) was 13 per firm and the maximum number was 388.

[86] These "fixed' inputs may only be fixed in the short run. Or we may simply decide to allow outputs and intermediate inputs to be variable and condition on an x vector of primary inputs.

[87] The concept of this function is due to Samuelson (1953).

[88] Let $x \geq 0_N$ and $p >> 0_M$. Then by (ii), there exists y_x such that $(y_x,x)\in S$. Define the closed and bounded set $B(x,p) \equiv \{y: y \leq b(x)1_M;\ p^Ty \geq p^Ty_x\}$ where $b(x) > 0$ is an upper bound on all possible net output vectors that can be produced by the technology if the vector of fixed inputs x is available to the producer. It can be seen that the constraint $(y,x)\in S$ in (213) can be replaced by the constraint $(y,x)\in S\cap B(x,p)$. Using (ii), $S\cap B(x,p)$ is a closed and bounded set so that the maximum in (213) will exist.

show that the profit function $\pi(p,x)$ satisfies some regularity conditions without assuming much about the technology set S.

Theorem 9: McFadden (1966) (1978), Gorman (1968), Diewert (1973): Suppose the technology set S satisfies the minimal regularity conditions (i) and (ii) above. Then the variable profit function $\pi(p,x)$ defined by (213) has the following properties with respect to p for each $x \geq 0_N$:

Property 1: $\pi(p,x)$ is *positively linearly homogeneous in p* for each fixed $x \geq 0_N$; i.e.,

(214) $\pi(\lambda p,x) = \lambda\pi(p,x)$ for all $\lambda > 0$, $p >> 0_M$ and $x \geq 0_N$.

Property 2: $\pi(p,x)$ is a *convex function of p* for each $x \geq 0_N$; i.e.,

(215) $x \geq 0_N$, $p^i >> 0_M$, $i = 1,2$; $0 < \lambda < 1$ implies
$\pi(\lambda p^1+(1-\lambda)p^2,x) \leq \lambda\pi(p^1,x) + (1-\lambda)\pi(p^2,x)$.

Problem

23. Prove Theorem 9. *Hint*: Properties 1 and 2 above for $\pi(p,x)$ are analogues to Properties 2 and 4 for the cost function C(y,p) in Theorem 1 above and can be proven in the same manner.

We now ask whether a knowledge of the profit function $\pi(p,x)$ is sufficient to determine the underlying technology set S. As was the case in section 3 above, the answer to this question is *yes*, but with some qualifications.

To see how to use a given profit function $\pi(p,x)$ can be used to determine the technology set that generated it, pick an arbitrary vector of fixed inputs $x \geq 0_N$ and an arbitrary vector of positive prices, $p^1 >> 0_M$. Now use the given profit function π to define the following isoprofit surface: $\{y: p^{1T}y = \pi(p^1,x)\}$. This isoprofit surface must be tangent to the set of net output combinations y that are feasible, given that the vector of fixed inputs x is available to the firm, which is the conditional on x production possibilities set, $S(x) \equiv \{x: (y,x)\in S\}$. It can be seen that this isoprofit surface and the set lying below it must contain the set S(x); i.e., the following *halfspace* $M(x,p^1)$, contains S(x):

(216) $M(x,p^1) \equiv \{y: p^{1T}y \leq \pi(p^1,x)\}$.

Pick another positive vector of prices, $p^2 >> 0_M$ and it can be seen, repeating the above argument, that the halfspace $M(x,p^2) \equiv \{y: p^{2T}y \leq \pi(p^2,x)\}$ must also contain the conditional on x production possibilities set S(x). Thus S(x) must belong to the intersection of the two halfspaces $M(x,p^1)$ and $M(x,p^2)$. Continuing to argue along these lines, it can be seen that S(x) must be contained in the following set, which is the intersection over all $p >> 0_M$ of all of the supporting halfspaces to S(x):

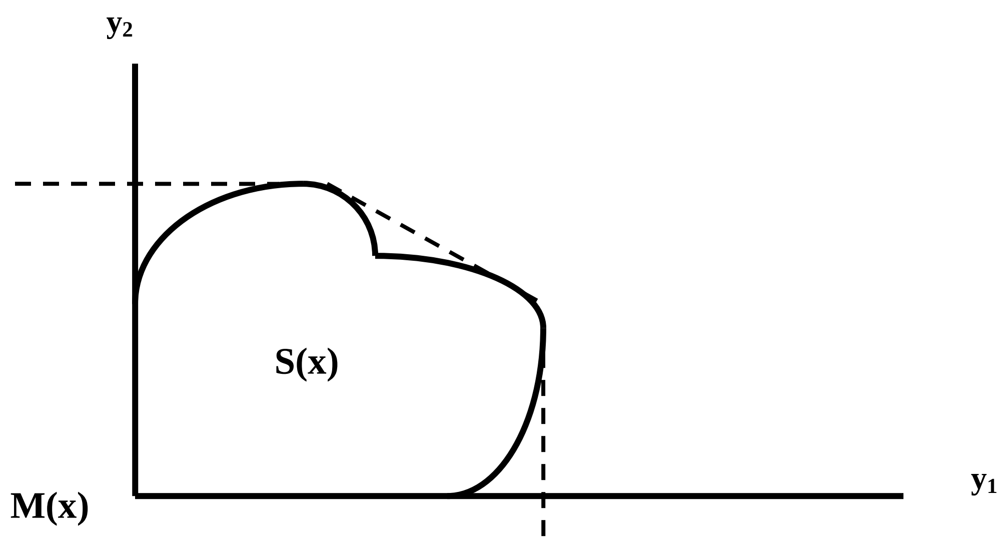
(217) $M(x) \equiv \bigcap_{p>>0_M} M(x,p)$.

Note that M(x) is defined using just the given profit function, $\pi(p,x)$. Note also that since each of the sets in the intersection, M(x,p), is a convex set, then M(x) is also a convex set. Since S(x) is a subset of each M(x,p), it must be the case that S(x) is also a subset of M(x); i.e., we have

(218) $S(x) \subset M(x)$.

Is it the case that S(x) is equal to M(x)? In general, the answer is *no*; M(x) forms an *outer approximation* to the true conditional production possibilities set S(x). Suppose that that there are only two outputs and for a given input vector x, the output production possibilities set is the heart shaped region in Figure 2. The boundary of the set M(x) partly coincides with the boundary of S(x) but it encloses a bigger set: the backward bending parts of the true production frontier are replaced by the dashed lines that are parallel to the y_1 axis and the y_2 axis and the inward bending part of the true production frontier is replaced by the dashed line that is tangent to the two regions where the boundary of M(x) coincides with the boundary of S(x). However, if the producer is a price taker in the two output markets, then it can be seen that *we will never observe the producer's nonconvex or backward bending parts of the production frontier*.

Figure 2: The Geometry of the Two Output Maximization Problem



What are conditions on the technology set S (and hence on the conditional technology sets S(x)) that will ensure that the outer approximation sets M(x), constructed using the variable profit function $\pi(p,x)$, will equal the true technology sets S(x)? It can be seen that the following two conditions on S (in addition to the minimal regularity conditions (i) and (ii)) are the required conditions:

(219) For every $x \geq 0_N$, the set $S(x) \equiv \{x: (y,x)\in S\}$ has the following *free disposal property*: $y^1 \in S(x)$, $y^2 \leq y^1$ implies $y^2 \in S(x)$;

(220) For every $x \geq 0_N$, the set $S(x) \equiv \{y: (y,x)\in S\}$ is convex.[89]

Conditions (219) and (220) are the conditions on the technology set S that are counterparts to the two regularity conditions of nondecreasingness and quasiconcavity[90] that were made on the production function, f(x), in section 3 above in order to obtain a duality between cost and production functions. If the firm is behaving as a price taker in variable commodity markets, it can be seen that it is not restrictive from an empirical point of view to assume that S satisfies conditions (219) and (220), just as it was not restrictive to assume that the production function was nondecreasing and quasiconcave in the context of the producer's (competitive) cost minimization problem studied earlier.

The next result provides a counterpart to Shephard's Lemma, Theorem 5 in section 4 above.

Theorem 10: Hotelling's (1932; 594) *Lemma*:[91] If the profit function $\pi(p,x)$ satisfies the properties listed in Theorem 9 above and in addition is once differentiable with respect to the components of the variable commodity prices at the point (p^*,x^*) where $x^* \geq 0_N$ and $p^* >> 0_M$, then

(221) $y^* = \nabla_p\pi(p^*,x^*)$

where $\nabla_p\pi(p^*,x^*)$ is the vector of first order partial derivatives of variable profit with respect to variable commodity prices and y^* is any solution to the profit maximization problem

(222) $\max_y \{p^{*T}y: (y,x^*)\in S\} \equiv \pi(p^*,x^*)$.

Under these differentiability hypotheses, it turns out that the y^* solution to (222) is unique.

Proof: Let y^* be any solution to the profit maximization problem (222). Since y^* is feasible for the profit maximization problem when the variable commodity price vector is changed to an arbitrary $p >> 0_M$, it follows that

(223) $p^Ty^* \leq \pi(p,x^*)$ for every $p >> 0_M$.

Since y^* is a solution to the profit maximization problem (22) when $p = p^*$, we must have

[89] If N = 1 so that there is only one fixed input, then given a producible net output vector $y\in R^M$, we can define the (fixed) *input requirements function* that corresponds to the technology set S as $g(y) \equiv \min_x \{x: (y,x)\in S\}$. In this case, condition (220) becomes the following condition: the input requirements function g(y) is *quasiconvex* in y. For additional material on this one fixed input model, see Diewert (1974b).

[90] Recall conditions (11) and (12) in section 3.

[91] See also Gorman (1968) and Diewert (1974a, 137).

(224) $p^{*T}y^* = \pi(p^*,x^*)$.

But (223) and (224) imply that the function of M variables, $g(p) \equiv p^Ty^* - \pi(p,x^*)$ is nonpositive for all $p >> 0_M$ with $g(p) = 0$. Hence, g(p) attains a global maximum at $p = p^*$ and since g(p) is differentiable with respect to the variable commodity prices p at this point, the following first order necessary conditions for a maximum must hold at this point:

(225) $\nabla_p\, g(p^*) = y^* - \nabla_p\pi(p^*,x^*) = 0_M$.

Now note that (225) is equivalent to (221). If y^{**} is any other solution to the profit maximization problem (222), then repeat the above argument to show that $y^{**} = \nabla_p\pi(p^*,x^*)$ which in turn is equal to y^*. Q.E.D.

Hotelling's Lemma may be used in order to derive systems of variable commodity output supply and input demand functions just as we used Shephard's Lemma to generate systems of cost minimizing input demand functions; for examples of this use of Hotelling's Lemma, see Diewert (1974a; 137-139) and Sections 17-19 below.

If we are willing to make additional assumptions about the underlying firm production possibilities set S, then we can deduce that $\pi(p,x)$ satisfies some additional properties. One such additional property is the following one: S is subject to the *free disposal of fixed inputs* if it has the following property:

(226) $x^2 > x^1 \geq 0_N$ and $(y,x^1)\in S$ implies $(y,x^2)\in S$.

The above property means if the vector of fixed inputs x^1 is sufficient to produce the vector of variable inputs and outputs y and if we have at our disposal a bigger vector of fixed inputs x^2, then y is still producible by the technology that is represented by the set S.

Theorem 11:[92] Suppose the technology set S satisfies the weak regularity conditions (i) and (ii) above. (a) If in addition, S has the following property:[93]

(227) For every $x \geq 0_N$, $(0_M,x)\in S$;

then for every $p >> 0_M$ and $x \geq 0_N$, $\pi(p,x) \geq 0$; i.e., the variable profit function is *nonnegative* if (227) holds.
(b) If S is a convex set, then for each $p >> 0_M$, then $\pi(p,x)$ is a *concave function* of x over the set $\Omega \equiv \{x: x \geq 0_N\}$.
(c) If S is a cone so that the technology is subject to constant returns to scale, then $\pi(p,x)$ is (positively) *homogeneous of degree one* in the components of x.

---

[92] The results in this Theorem are essentially due to Samuelson (1953; 20), Gorman (1968), McFadden (1968) and Diewert (1973) (1974a; 136) but they are packaged in a somewhat different form in this chapter.
[93] This property says that the technology can always produce no variable outputs and utilize no variable inputs given any vector of fixed inputs x.

(d) If S is subject to the free disposal of fixed inputs, property (226), then

(228) $p >> 0$, $x^2 > x^1 \geq 0_N$ implies $\pi(p,x^2) \geq \pi(p,x^1)$;

i.e., $\pi(p,x)$ is *nondecreasing* in the components of x.

Proof of (a): Let $p >> 0_M$ and $x \geq 0_N$. Then

(229) $\pi(p,x) \equiv \max_y \{p^Ty: (y,x)\in S\}$
$\geq p^T 0_M$ since by (227), $(0_M,x)\in S$ and hence is feasible for the problem
$= 0$.

Proof of (b): Let $p >> 0_M$, $x^1 \geq 0_N$, $x^2 \geq 0_N$ and $0 < \lambda < 1$. Then

(230) $\pi(p,x^1) \equiv \max_y \{p^Ty: (y,x^1)\in S\}$
$= p^Ty^1$ where $(y^1,x^1)\in S$;

(231) $\pi(p,x^2) \equiv \max_y \{p^Ty: (y,x^2)\in S\}$
$= p^Ty^2$ where $(y^2,x^2)\in S$.

Since S is assumed to be a convex set, we have

(232) $\lambda(y^1,x^1) + (1-\lambda)(y^2,x^2) = [\lambda y^1 + (1-\lambda)y^2, \lambda x^1 + (1-\lambda)x^2]\in S$.

Using the definition of π, we have:

(233) $\pi(p,\lambda x^1 + (1-\lambda)x^2) \equiv \max_y \{p^Ty: (y,\lambda x^1 + (1-\lambda)x^2)\in S\}$
$\geq p^T[\lambda y^1 + (1-\lambda)y^2]$ since by (232), $\lambda y^1 + (1-\lambda)y^2$ is feasible for the problem
$= \lambda p^Ty^1 + (1-\lambda)p^Ty^2$
$= \lambda\pi(p,x^1) + (1-\lambda)\pi(p,x^2)$ using (230) and (231).

Proof of (c): Let $p >> 0_M$, $x^* \geq 0_N$ and $\lambda > 0$. Then

(234) $\pi(p,x^*) \equiv \max_y \{p^Ty: (y,x^*)\in S\}$
$= p^Ty^*$ where $(y^*,x^*)\in S$.

Since S is a cone and since $(y^*,x^*)\in S$, then we have $(\lambda y^*,\lambda x^*)\in S$ as well. Hence, using a feasibility argument:

(235) $\pi(p,\lambda x^*) \equiv \max_y \{p^Ty: (y,\lambda x^*)\in S\}$
$\geq p^T\lambda y^*$ since $(\lambda y^*,\lambda x^*)\in S$ and hence is feasible for the problem
$= \lambda p^Ty^*$.

Now *suppose* that the strict inequality in (235) held so that

$$(236)\ \pi(p,\lambda x^*) \equiv \max_y \{p^Ty: (y,\lambda x^*)\in S\}$$
$$= p^Ty^{**} \qquad \text{where } (y^{**},\lambda x^*)\in S$$
$$> \lambda p^Ty^*.$$

Since S is a cone, $\lambda > 0$, and $(y^{**},\lambda x^*)\in S$, then we have $(\lambda^{-1}y^{**},x^*)\in S$ as well. Thus $\lambda^{-1}y^{**}$ is feasible for the maximization problem (234) that defined $\pi(p,x^*)$ and so

$$(237)\ p^Ty^* = \max_y \{p^Ty: (y,x^*)\in S\} \qquad \text{using (234)}$$
$$\geq p^T\lambda^{-1}y^{**} \qquad \text{since } \lambda^{-1}y^{**} \text{ is feasible for the problem}$$
$$= \lambda^{-1}p^Ty^{**}$$

or since $\lambda > 0$, (237) is equivalent to

$$(238)\ \lambda p^Ty^* \geq p^Ty^{**} > \lambda p^Ty^* \qquad \text{using (236).}$$

But (238) implies that $\lambda p^Ty^* > \lambda p^Ty^*$, which is impossible and hence our *supposition* is false and the desired result follows.

Proof of (d): Let $p >> 0_M$, $x^2 > x^1 \geq 0_N$. Using the definition of $\pi(p,x^1)$, we have

$$(239)\ \pi(p,x^1) \equiv \max_y \{p^Ty: (y,x^1)\in S\}$$
$$= p^Ty^1 \qquad \text{where } (y^1,x^1)\in S.$$

Using the free disposal property (228) for S, since $(y^1,x^1)\in S$ and $x^2 > x^1$, we have

$$(240)\ (y^1,x^2)\in S.$$

Using the definition of $\pi(p,x^2)$, we have

$$(241)\ \pi(p,x^2) \equiv \max_y \{p^Ty: (y,x^2)\in S\}$$
$$\geq p^Ty^1 \qquad \text{since by (240), } (y^1,x^2) \text{ is feasible}$$
$$= \pi(p,x^1) \qquad \text{using (239).}$$

Q.E.D.

If the technology set S satisfies the minimal regularity conditions (i) and (ii) plus all of the additional conditions that are listed in Theorem 11 above (we shall call such a technology set a *regular technology set*), then the associated variable profit function $\pi(p,x)$ will have *all* of the regularity conditions with respect to its fixed input vector x that a nonnegative, nondecreasing, concave and linearly homogeneous production function f(x) possesses with respect to its input vector x.

Hotelling's Lemma enabled us to interpret the vector of first order partial derivatives of the variable profit function with respect to the components of the variable commodity price vector p, $\nabla_p\pi(p,x)$, as the producer's vector of variable profit maximizing output supply (and the negative of variable input demand) functions, y(p,x), provided that the derivatives existed. If the first order partial derivatives of the variable profit function

$\pi(p,x)$ with respect to the components of the fixed input vector x exist, then this vector of derivatives, $\nabla_x\pi(p,x)$, can also be given an economic interpretation as a vector of *shadow prices* or imputed contributions to profit of adding marginal units of fixed inputs. The following result also shows that these derivatives can be interpreted as competitive input prices for the "fixed" factors if they are allowed to become variable.

Theorem 12; *Samuelson's Lemma*:[94] Suppose the technology set S satisfies the minimal regularity assumptions (i) and (ii) above and in addition is a convex set. Suppose in addition that $p^* >> 0_M$, $x^* \geq 0_N$ and that the vector of derivatives, $\nabla_x\pi(p^*,x^*) \equiv w^*$, exists. Then x^* is a solution to the following *long run profit maximization problem* that allows the "fixed" inputs x to be variable:

(242) $\max_x \{\pi(p^*,x) - w^{*T}x: x \geq 0_N\}$.

Proof: Part (b) of Theorem 11 above implies that $\pi(p^*,x)$ is a concave function of x over the set $\Omega \equiv \{x: x \geq 0_N\}$. The function $-w^{*T}x$ is linear in x and hence is also a concave function of x over Ω. Hence f(x) defined for $x \geq 0_N$ as

(243) $f(x) \equiv \pi(p^*,x) - w^{*T}x$

is also a concave function in x over the set Ω. Since $x^* \geq 0_N$, $x^* \in \Omega$. Hence using the fact that a differentiable concave function has a Taylor series approximation that provides an upper bound to the function around any point x^* where the function is differentiable, we have the following inequality:

(244) $f(x) \leq f(x^*) + \nabla_x f(x^*)^T(x - x^*)$ for all $x \geq 0_N$
$= \pi(p^*,x^*) - w^{*T}x^* + 0_N^T(x - x^*)$ since $\nabla_x f(x^*) = \nabla_x\pi(p^*,x^*) - w^* = 0_N$
$= \pi(p^*,x^*) - w^{*T}x^*$.

But (243) and (244) show that x^* solves the profit maximization problem (242). Q.E.D.

Corollary: If in addition to the above assumptions, $\pi(p,x)$ is differentiable with respect to the components of p at the point (p^*,x^*), so that $y^* \equiv \nabla_p\pi(p^*,x^*)$ exists, then (y^*,x^*) solves the following *long run profit maximization problem*:

(245) $\Pi(p^*,w^*) \equiv \max_{y,x} \{p^{*T}y - w^{*T}x: (y,x)\in S\}$.

Proof: Using Hotelling's Lemma, we know that y^* solves the following variable profit maximization problem:

(246) $\pi(p^*,x^*) \equiv \max_y \{p^{*T}y: (y,x^*)\in S\} = p^{*T}y^*$.

[94] Samuelson's National Product function, N(p,v), is the counterpart to our $\pi(p,x)$ where his v is a vector of primary inputs. Samuelson (1953; 10) derived the equations $w = \nabla_v N(p,v)$. Our proof follows that of Diewert (1974a; 140).

Now look at the long run profit maximization problem defined by (245):

$$(247)\ \Pi(p^*,w^*) \equiv \max_{y,x}\{ p^{*T}y - w^{*T}x: (y,x)\in S\}$$
$$= \max_x\{\max_y [p^{*T}y: (y,x)\in S] - w^{*T}x\} \quad \text{where we have rewritten the maximization problem as a two stage maximization problem}$$
$$= \max_x [\pi(p^*,x) - w^{*T}x] \quad \text{using the definition of } \pi(p^*,x)$$
$$= \pi(p^*,x^*) - w^{*T}x^* \quad \text{using Theorem 12.}$$

Hence with $x = x^*$ being an x solution to (247), we must have

$$(248)\ \Pi(p^*,w^*) \equiv \max_{y,x} \{p^{*T}y - w^{*T}x: (y,x)\in S\}$$
$$= \max_y\{[p^{*T}y: (y,x^*)\in S] - w^{*T}x^*\} \quad \text{letting } x = x^*$$
$$= p^{*T}y^* - w^{*T}x^* \quad \text{using (246).} \qquad \text{Q.E.D.}$$

Hotelling's Lemma and Samuelson's Lemma can be used as a convenient method for obtaining econometric estimating equations for determining the parameters that characterize a producer's technology set S. Assuming that S satisfies the minimal regularity conditions on S, we need only postulate a differentiable functional form for the producer's variable profit function, $\pi(p,x)$, that is linearly homogeneous and convex in p. Suppose that we have collected data on the fixed input vectors used by the production unit in period t, x^t, and the net supply vectors for variable commodities produced in period t, y^t, for $t = 1,\ldots,T$ time periods as well as the corresponding variable commodity price vectors p^t. Then the following MT equations can be used in order to estimate the unknown parameters in $\pi(p,x)$:

$$(249)\ y^t = \nabla_p\pi(p^t,x^t) + u^t\ ; \qquad t = 1,\ldots,T$$

where u^t is a vector of errors. If in addition, S is a convex set and it can be assumed that the production unit is optimizing with respect to its vector of "fixed" inputs in each period, where it faces the "fixed" input price vector w^t in period t, then the following N equations can be added to (249) as additional estimating equations:

$$(250)\ w^t = \nabla_x\pi(p^t,x^t) + v^t\ ; \qquad t = 1,\ldots,T$$

where v^t is a vector of errors.[95] We will look at some specific functional forms for $\pi(p,x)$ and their econometric estimating equations in the final sections of this chapter.

16. The Comparative Statics Properties of Net Supply and Fixed Input Demand Functions

---

[95] If in addition, the technology set S is subject to constant returns to scale and the data reflect this fact by "adding up" (so that $p^{tT}y^t = w^{tT}x^t$ for $t = 1,\ldots,T$), then the error vectors u^t and v^t in (249) and (250) cannot be statistically independent. Hence, under these circumstances, one of the M+N equations in (249) and (250) must be dropped from the system of estimating equations.

From Theorem 11 above, we know that the firm's variable profit function $\pi(p,x)$ is convex and linearly homogeneous in the components of the vector of variable commodity prices p for each fixed input vector x. Thus if $\pi(p,x)$ is twice continuously differentiable with respect to the components of p at some point (p,x), then using Hotelling's Lemma, we can prove the following counterpart to Theorem 7 for the cost function.

Theorem 13: Hotelling (1932; 597), Hicks (1946; 321), Samuelson (1953; 10), Diewert (1974a; 142-146): Suppose the variable profit function $\pi(p,x)$ is linearly homogeneous and convex in p and in addition is twice continuously differentiable with respect to the components of p at some point, (p,x). Then the *system of variable profit maximizing net supply functions*, $y(p,x) \equiv [y_1(p,x),\ldots,y_M(p,x)]^T$, exists at this point and these net supply functions are once continuously differentiable. Form the M by M matrix of net supply derivatives with respect to variable commodity prices, $B \equiv [\partial y_m(p,x)/\partial p_k]$, which has mk element equal to $\partial y_m(p,x)/\partial p_k$. Then the matrix B has the following properties:

(251) $B = B^T$ so that $\partial y_m(p,x)/\partial p_k = \partial y_k(p,x)/\partial p_m$ for all $m \neq k$;[96]
(252) B is positive semidefinite and
(253) $Bp = 0_M$.

Proof: Hotelling's Lemma implies that the firm's system of variable profit maximizing net supply functions, $y(p,x) \equiv [y_1(p,x),\ldots,y_M(p,x)]^T$, exists and is equal to

(254) $y(p,x) = \nabla_p \pi(p,x)$.

Differentiating both sides of (254) with respect to the components of p gives us

(255) $B \equiv [\partial y_m(p,x)/\partial p_k] = \nabla^2_{pp}\pi(p,x)$.

Property (251) follows from Young's Theorem in calculus. Property (252) follows from (255) and the fact that $\pi(p,x)$ is convex and twice differentiable in p and hence the matrix of second order partial derivatives $\nabla^2_{pp}\pi(p,x)$ must be positive semidefinite. Finally, property (253) follows from the fact that the profit function is linearly homogeneous in p and hence, using Part 2 of Euler's Theorem on homogeneous functions, (253) holds. Q.E.D.

Note that property (252) implies the following properties on the net supply functions:

(256) $\partial y_m(p,x)/\partial p_m \geq 0$ for $m = 1,\ldots,M$.

Property (256) means that output supply curves cannot be downward sloping. However, if variable commodity m is an input, then $y_m(p,x)$ is negative. If we define the positive input demand function as

(257) $d_m(p,x) \equiv - y_m(p,x) \geq 0$,

[96] These are the Hotelling (1932; 549) and Hicks (1946; 321) symmetry restrictions on supply functions.

then the restriction (256) translates into $\partial d_m(p,x)/\partial p_m \leq 0$, which means that variable input demand curves cannot be upward sloping.

Obviously, if the technology set is a convex cone, then the firm's competitive fixed input price functions (or inverse demand functions), $w(p,x) \equiv \nabla_x \pi(p,x)$, will satisfy properties analogous to the properties of cost minimizing input demand functions in Theorem 7.

Theorem 14: Samuelson (1953; 10), Diewert (1974a; 144-146): Suppose that the production unit's technology set S is regular. Define the variable profit function $\pi(p,x)$ by (213). Suppose that $\pi(p,x)$ is twice continuously differentiable with respect to the components of x at some point (p,x) where $p >> 0_M$ and $x \geq 0_N$. Then the *system of input price functions*, $w(p,x) \equiv [w_1(p,x),\ldots,w_N(p,x)]^T$, exists at this point[97] and these input price functions are once continuously differentiable. Form the N by N matrix of input price derivatives with respect to the "fixed" inputs, $C \equiv [\partial w_i(p,x)/\partial x_k]$, which has ik element equal to $\partial w_i(p,x)/\partial x_k$. Then the matrix C has the following properties:

(258) $C = C^T$ so that $\partial w_i(p,x)/\partial x_k = \partial w_k(p,x)/\partial x_i$ for all $i \neq k$;
(259) C is negative semidefinite and
(260) $Cx = 0_N$.

Proof: Using Samuelson's Lemma, the firm's system of fixed input price functions, $w(p,x) \equiv [w_1(p,x),\ldots,w_N(p,x)]^T$, exists and is equal to

(261) $w(p,x) = \nabla_x \pi(p,x)$.

Differentiating both sides of (261) with respect to the components of x gives us

(262) $C \equiv [\partial w_i(p,x)/\partial x_k] = \nabla^2_{xx} \pi(p,x)$.

Now property (258) follows from Young's Theorem in calculus. Property (259) follows from (262) and the fact that $\pi(p,x)$ is concave in x.[98] Finally, property (260) follows from the fact that the profit function is linearly homogeneous in x[99] and hence, using Part 2 of Euler's Theorem on homogeneous functions, (260) holds Q.E.D.

Note that property (259) implies the following properties on the fixed input price functions:

---

[97] The assumption that S is regular implies that S has the free disposal property in fixed inputs property (226), which implies by part (d) of Theorem 11 that $\pi(p,x)$ is nondecreasing in x and this in turn implies that $w(p,x) \equiv \nabla_x \pi(p,x)$ is nonnegative.

[98] The assumption that S is regular implies that S is a convex set and this in turn implies that $\pi(p,x)$ is concave in x. Concavity in x plus our differentiability assumption implies that $\nabla^2_{xx}\pi(p,x)$ is negative semidefinite.

[99] The assumption that S is regular implies that S is a cone and this in turn implies that $\pi(p,x)$ is linearly homogeneous in x.

(263) $\partial w_n(p,x)/\partial x_n \le 0$; n = 1,…,N.

Property (263) means that the inverse fixed input demand curves cannot be upward sloping.

If the firm's production possibilities set S is regular and if the corresponding variable profit function $\pi(p,x)$ is twice continuously differentiable with respect to all of its variables, then there will be additional restrictions on the derivatives of the variable net output supply functions $y(p,x) = \nabla_p\pi(p,x)$ and on the derivatives of the fixed input price functions $w(p,x) = \nabla_x\pi(p,x)$. Define the M by N *matrix of derivatives of the net output supply functions* y(p,x) *with respect to the components of the vector of fixed inputs* x as follows:

(264) $D \equiv [\partial y_m(p,x)/\partial x_n] = \nabla^2_{px}\pi(p,x)$; m = 1,…,M; n = 1,…,N,

where the equalities in (264) follow by differentiating both sides of the Hotelling's Lemma relations, $y(p,x) = \nabla_p\pi(p,x)$, with respect to the components of x. Similarly, define the N by M *matrix of derivatives of the fixed input price functions* w(p,x*) with respect to the components of the vector of variable commodity prices p* as follows:

(265) $E \equiv [\partial w_n(p,x)/\partial p_m] = \nabla^2_{xp}\pi(p,x)$; n = 1,…,N; m = 1,…,M,

where the equalities in (265) follows by differentiating both sides of the Samuelson's Lemma relations, $w(p,x) = \nabla_x\pi(p,x)$, with respect to the components of p.

Theorem 15: Samuelson (1953; 10), Diewert (1974a; 144-146): Suppose that the production unit's technology set S is regular. Define the variable profit function $\pi(p,x)$ by (213). Suppose that $\pi(p,x)$ is twice continuously differentiable with respect to the components of x at some point (p,x) where $p >> 0_M$ and $x \ge 0_N$ and define the matrices of derivatives D and E by (264) and (265) respectively. Then these matrices have the following properties:

(266) $D = E^T$ so that $\partial y_m(p,x)/\partial x_n = \partial w_n(p,x)/\partial x_m$ for m = 1,…,M and n = 1,…,N;
(267) $w(p,x) = Ep \ge 0_N$;
(268) $y(p,x) = Dx$.

Proof: The symmetry restrictions (266) follow from definitions (264) and (265) and Young's Theorem in calculus.

Since $\pi(p,x)$ is linearly homogeneous in the components of p, we have

(269) $\pi(\lambda p,x) = \lambda\pi(p,x)$ for all $\lambda > 0$.

Partially differentiate both sides of (269) with respect to x_n and we obtain:

(270) $\partial\pi(\lambda p,x)/\partial x_n = \lambda\partial\pi(p,x)/\partial x_n$ for all $\lambda > 0$ and $n = 1,\ldots,N$.

But (270) implies that the functions $w_n(p,x) \equiv \partial\pi(p,x)/\partial x_n$ are homogeneous of degree one in p. Hence, we can apply Part 1 of Euler's Theorem on homogeneous functions to these functions $w_n(p,x)$ and conclude that

(271) $w_n(p,x) = \sum_{m=1}^{M} [\partial w_n(p,x)/\partial p_m]p_m$; $\qquad n = 1,\ldots,N$.

But equations (271) are equivalent to the equations in (267). The inequality in (267) follows from $w(p,x) = \nabla_x\pi(p,x) \geq 0_N$, which in turn follows from the fact that regularity of S implies that $\pi(p,x)$ is nondecreasing in the components of x.

Since S is regular, part (c) of Theorem 11 implies that $\pi(p,x)$ is linearly homogeneous in x, so that

(272) $\pi(p,\lambda x) = \lambda\pi(p,x)$ for all $\lambda > 0$.

Partially differentiate both sides of (272) with respect to p_m and we obtain:

(273) $\partial\pi(p,\lambda x)/\partial p_m = \lambda\partial\pi(p,x)/\partial p_m$ for all $\lambda > 0$ and $m = 1,\ldots,M$.

But (273) implies that the functions $y_m(p,x) \equiv \partial\pi(p,x)/\partial p_m$ are homogeneous of degree one in x. Hence, we can apply Part 1 of Euler's Theorem on homogeneous functions to these functions $y_m(p,x)$ and conclude that

(274) $y_m(p,x) = \sum_{n=1}^{N} [\partial y_m(p,x)/\partial x_n]x_n$; $\qquad m = 1,\ldots,M$.

But equations (274) are equivalent to equations (268). Q.E.D.

Following up on the pioneering work of Samuelson (1953), Diewert and Woodland (1977; 383-390) developed additional comparative statics properties for a consolidated production sector consisting of a finite number of constant returns to scale production units. For additional applications of the National Product Function to the theory of international trade, see Kohli (1978) (1991), Dixit and Norman (1980), Woodland (1982) and Feenstra (2004).

Problems

24. Under the hypotheses of Theorem 15, show that y(p,x) and w(p,x) satisfy the following equation:

(i) $p^T y(p,x) = x^T w(p,x)$.

25. Let S be a technology set that satisfies the minimal regularity assumptions and let $\pi(p,x)$ be the corresponding differentiable variable profit function defined by (213).

Variable commodities m and k (where $m \neq k$) are said to be *substitutes* if (i) below holds, *unrelated* if (ii) below holds and *complements* if (iii) below holds:

(i) $\partial y_m(p,x)/\partial p_k < 0$;
(ii) $\partial y_m(p,x)/\partial p_k = 0$;
(iii) $\partial y_m(p,x)/\partial p_k > 0$.

(a) If the number of variable commodities $M = 2$, then show that the two variable commodities cannot be complements.
(b) If $M = 2$ and the two variable commodities are unrelated, then show that:

(iv) $\partial y_1(p,x)/\partial p_1 = \partial y_2(p,x)/\partial p_2 = 0$.

(c) If $M = 3$, then show that at most one pair of variable commodities can be complements.[100]

26. Let S be a regular technology set and let $\pi(p,x)$ be the corresponding twice continuously differentiable variable profit function defined by (213). Variable commodities m and fixed input n are said to be *normal* if (i) below holds, *unrelated* if (ii) below holds and *inferior* if (iii) below holds (we assume $p >> 0_M$ and $x >> 0_N$):

(i) $\partial y_m(p,x)/\partial x_n = \partial w_n(p,x)/\partial p_m > 0$;
(ii) $\partial y_m(p,x)/\partial x_n = \partial w_n(p,x)/\partial p_m = 0$;
(iii) $\partial y_m(p,x)/\partial x_n = \partial w_n(p,x)/\partial p_m < 0$.

(a) If $w_n(p,x) > 0$, then there exists at least one variable commodity m such that commodity m and fixed input n are normal.
(b) If $w_n(p,x) \geq 0$, then there exists at least one variable commodity m such that commodity m and fixed input n are either normal or unrelated.
(c) If $y_m(p,x) > 0$, then there exists at least one fixed input n such that commodity m and fixed input n are normal.
(d) If $y_m(p,x) < 0$, then there exists at least one fixed input n such that commodity m and fixed input n are inferior.

In the following three sections, we will look at the properties of some specific functional forms for a variable profit function. We will assume that these profit functions are dual to a regular technology.

17. The Translog Variable Profit Function

Assume that the log of the variable profit function for a regular technology, $\ln\pi(p,x)$, has the following *translog* functional form:[101]

---

[100] This type of argument (that substitutability tends to be more predominant than complementarity) is again due to Hicks (1946; 322-323) but we have not followed his terminology exactly.

(275) $\ln\pi(p,x) \equiv a_0 + \sum_{m=1}^{M} a_m \ln p_m + (1/2) \sum_{m=1}^{M}\sum_{k=1}^{M} a_{mk} \ln p_m \ln p_k$
$+ \sum_{n=1}^{N} b_n \ln x_n + (1/2)\sum_{n=1}^{N}\sum_{i=1}^{N} b_{ni} \ln x_n \ln x_i + \sum_{m=1}^{M}\sum_{n=1}^{N} c_{mn} \ln p_m \ln x_n$;

The coefficients must satisfy the following restrictions in order for $\pi(p,x)$ to be linearly homogeneous in the components of p as well as the components of x:[102]

(276) $\sum_{m=1}^{M} a_m = 1$;
(277) $\sum_{n=1}^{N} b_n = 1$;
(278) $a_{mk} = a_{km}$ for all k,m ;
(279) $b_{ni} = b_{in}$ for all n,i ;
(280) $\sum_{k=1}^{M} a_{mk} = 0$ for m = 1,...,M;
(281) $\sum_{i=1}^{N} b_{ni} = 0$ for n = 1,...,N;
(282) $\sum_{n=1}^{N} c_{mn} = 0$ for m = 1,...,M;
(283) $\sum_{m=1}^{M} c_{mn} = 0$ for n = 1,..., N.

If some of the variable outputs are actually inputs, then the domain of definition of p and x needs to be restricted to p and x such that $\pi(p,x) > 0$, since we cannot take the logarithm of a non-positive number. The proof that the translog profit function defined by (275)-(283) is linearly homogeneous in p follows our earlier proof that the translog unit cost function c(p) defined in section 9 was linearly homogeneous in p. The proof that $\pi(p,\lambda x) = \lambda\pi(p,x)$ for all $\lambda > 0$ follows in an analogous manner.

Note that using Hotelling's Lemma, we have $\partial\ln\pi(p,x)/\ln p_m = [p_m/\pi(p,x)]\partial\pi(p,x)/\partial p_m = [p_m/\pi(p,x)]y_m(p,x) \equiv s_m(p,x)$ where $y_m(p,x)$ is the profit maximizing conditional net supply function for net output m and $s_m(p,x)$ is the *share of net output m in total variable profits*. Thus differentiating the logarithm of $\pi(p,x)$ defined by (275) with respect to the logarithm of p_m leads to the following system of net variable output share equations:

(284) $s_m(p,x) = a_m + \sum_{k=1}^{M} a_{mk} \ln p_k + \sum_{n=1}^{N} c_{mn} \ln x_n$; m = 1,...,M.

Thus if we have data on the net outputs for period t, y^t, the corresponding net output prices $p^t >> 0_M$ and fixed inputs used in period t, $x^t >> 0_N$ by a production unit for t = 1,...,T, then we can form observed variable profits for period t, $\pi^t \equiv p^{tT}y^t > 0$ and the period t net variable output shares $s_m^t \equiv p_m^t y_m^t/\pi^t$ for m = 1,...,M and t = 1,...,T. A set of econometric estimating equations is the following very simple system of equations:

---

[101] This functional form was suggested by Diewert (1974a; 139) as a generalization of the translog functional form introduced by Christensen, Jorgenson and Lau (1971). Diewert (1974a; 139) indicated that this functional form was flexible for regular technologies. For applications of this functional form to international trade theory, see Kohli (1978) (1991). For applications to index number theory and the measurement of productivity, see Caves, Christensen and Diewert (1982a) (1982b), Diewert and Morrison (1986), Kohli (1990), Feenstra, Inklaar and Timmer (2015) and Inklaar and Diewert (2016).
[102] There are additional restrictions on the parameters which are necessary to ensure that $\pi(p,x)$ is convex in p and concave in x.

$$(285)\ \ln\pi^t = a_0 + \sum_{m=1}^{M} a_m \ln p_m^t + (1/2) \sum_{m=1}^{M}\sum_{k=1}^{M} a_{mk} \ln p_m^t \ln p_k^t + \sum_{n=1}^{N} b_n \ln x_n^t + (1/2)\sum_{n=1}^{N}\sum_{j=1}^{N} b_{nj} \ln x_n^t \ln x_j^t + \sum_{m=1}^{M}\sum_{n=1}^{N} c_{mn} \ln p_m^t \ln x_n^t + e_0^t; \quad t = 1,\ldots,T;$$

$$(286)\ s_m^t = a_m + \sum_{k=1}^{M} a_{mk} \ln p_k^t + \sum_{n=1}^{N} c_{mn} \ln x_n^t + e_m^t ; \quad m = 1,\ldots,M;\ t = 1,\ldots,T$$

where the e_m^t are error terms with 0 means for m = 0,1,...,M and t = 1,...,T. Note that these equations are linear in the unknown parameters. The cross equation symmetry restrictions, $a_{mk} = a_{km}$ for $1 \le m < k \le M$ could be imposed on the above equations or these conditions could be tested.[103]

Suppose now that we have reason to believe that the producer is optimizing with respect to the vector of "fixed" inputs x. Using Samuelsons's Lemma, we have $\partial\ln\pi(p,x)/\ln x_n = [x_n/\pi(p,x)]\partial\pi(p,x)/\partial x_n = [x_n/\pi(p,x)]w_n(p,x) \equiv S_m(p,x)$ where $w_n(p,x)$ is the profit maximizing inverse demand function for "fixed" input n and $S_n(p,x)$ is the *share of "fixed" input n in total "fixed" input cost*.[104] Thus differentiating the logarithm of $\pi(p,x)$ defined by (275) with respect to the logarithm of x_n leads to the following system of input cost share equations:

$$(287)\ S_n(p,x) = b_n + \sum_{j=1}^{N} b_{nj} \ln x_j + \sum_{m=1}^{M} c_{mn} \ln p_m ; \quad n = 1,\ldots,N.$$

Thus if we have data on "fixed" input prices for the T periods in addition to the already mentioned data, then we can form the observed cost shares for "fixed" input n in period t, $S_n^t \equiv w_n^t x_n^t/\pi^t$ for t = 1,...,T. Thus we can add the following set of estimating equations to the estimating equations defined by (285) and (286):

$$(288)\ S_n^t \equiv b_n + \sum_{j=1}^{N} b_{nj} \ln x_j^t + \sum_{m=1}^{M} c_{mn} \ln p_m^t + u_n^t ; \quad n = 1,\ldots,N;\ t = 1,\ldots,T$$

where the u_n^t are error terms with 0 means.[105]

The simplicity of the estimating equations given by (285), (286) and (288) means that it is relatively easy to estimate the translog variable profit function. However, there are two disadvantages associated with the translog functional form:

- Not all of the parameters of the translog $\pi(p,x)$ can be estimated unless equations (285) are included in the estimation procedure. But every parameter is included in each of these equations and this can lead to singularity problems if N + M is large and T is small;[106]

[103] Since the shares s_m^t sum to 1 over m for each t, the equations (286) cannot have independent error terms and hence one of the M equations in (286) should be dropped when estimating the unknown parameters.
[104] From Problem 24, we know that $p^T y(p,x) = x^T w(p,x) = \pi(p,x)$ since we have assumed that the underlying production possibilities set is regular.
[105] Since the shares S_n^t sum to one over n for each t, one of the N estimating equations in (287) should be dropped. Typically, the cross equation parameter restrictions defined by (278)-(283) would be imposed but in principle, they could be tested.
[106] This problem is perhaps not too serious; if equations (286) and (288) are estimated, then all of the parameters that appear in definition (275) can be identified except the parameter a_0. This parameter could

- It is not possible to impose the convexity in p and concavity in x property for the translog functional form without destroying the flexibility of the functional form.

Thus in the following two sections, we look at functional forms for a regular technology variable profit function where we can impose the correct concavity and convexity properties.

Another problem with the translog $\pi(p,x)$ defined by (275) is that this functional form does not allow for technical progress. This problem can be readily remedied: simply add the following terms to the right hand side of definition (275): $\alpha_0 t + \sum_{m=1}^{M} \alpha_m \, t\ln p_m + \sum_{n=1}^{N} \beta_n \, t\ln x_n$ where t is a scalar time variable and the new parameters α_m and β_n satisfy the additional restrictions $\sum_{m=1}^{M} \alpha_m = 0$ and $\sum_{n=1}^{N} \beta_n = 0$.[107] These restrictions will ensure that the resulting translog $\pi(p,x)$ is linearly homogeneous in p and x separately.[108]

18. The Normalized Quadratic Variable Profit Function

At this point, it will be useful to list the equations that a twice continuously differentiable functional form for a variable profit function $\pi(p,x)$ that is dual to a regular technology must satisfy in order to be a flexible functional form at the point $p^* >> 0_M$ and $x^* >> 0_N$. Let $\pi^*(p,x)$ be an arbitrary variable profit function that is dual to a regular technology set and suppose that $\pi^*(p,x)$ is twice continuously differentiable at (p^*,x^*). For $\pi(p,x)$ to be a flexible functional form, it must have enough free parameters so that it can provide a second order approximation to $\pi^*(p,x)$ at the point (p^*,x^*). Thus the candidate function π must have enough parameters so that it can satisfy the following $1 + M + N + (M+N)^2$ equations:

(289) $\pi(p^*,x^*) = \pi^*(p^*,x^*)$; 1 equation;
(290) $\nabla_p\pi(p^*,x^*) = \nabla_p\pi^*(p^*,x^*)$; M equations;
(291) $\nabla_x\pi(p^*,x^*) = \nabla_x\pi^*(p^*,x^*)$; N equations;
(292) $\nabla^2_{pp}\pi(p^*,x^*) = \nabla^2_{pp}\pi^*(p^*,x^*)$; M^2 equations;
(293) $\nabla^2_{xx}\pi(p^*,x^*) = \nabla^2_{xx}\pi^*(p^*,x^*)$; N^2 equations;
(294) $\nabla^2_{px}\pi(p^*,x^*) = \nabla^2_{px}\pi^*(p^*,x^*)$; MN equations;
(295) $\nabla^2_{xp}\pi(p^*,x^*) = \nabla^2_{xp}\pi^*(p^*,x^*)$; NM equations.

However, because $\pi(p,x)$ and $\pi^*(p,x)$ are both linearly homogeneous in p and x separately and both are assumed to be twice continuously differentiable at (p^*,x^*), not all of the equations in (289)-(295) are independent. Equation (289) is implied by the first part of Euler's Theorem on homogeneous functions and equations (290) or (291). Thus equation

be estimated in a second stage where equations (285) are used to solve for a_0 in terms of $\ln\pi^t$ and the fitted values from the first stage for the right hand side of equations (285) omitting the term a_0.

[107] This extension of the translog function GDP function to allow for technical progress is due to Kohli (1978) in a model with 4 outputs and 2 inputs. Feenstra (2004; 423) noted these restrictions in the general M outputs and N inputs model.

[108] More general specification of technical progress can be made using linear or quadratic splines in the time variable.

(289) can be dropped from the list of equations that $\pi(p,x)$ must satisfy since it will be satisfied if either (290) or (291) is satisfied. Since $p^{*T}\nabla_p\pi(p^*,x^*) = x^{*T}\nabla_x\pi(p^*,x^*)$ and $p^{*T}\nabla_p\pi^*(p^*,x^*) = x^{*T}\nabla_x\pi\,(p^*,x^*)$, any one of the M + N equations in (290) and (291) can also be dropped. Young's Theorem from calculus and the second part of Euler's Theorem on homogeneous functions imply that if the M(M−1)/2 equations in the upper triangle of equations (292) hold, then all M^2 equations in (292) will hold. Similarly, if the N(N−1)/2 equations in the upper triangle of equations (293) hold, then all N^2 equations in (293) will hold. Young's Theorem implies that if the MN equations in (294) hold, then the NM equations in (295) will also hold. Recall equations (254), (261), (267) and (268) in Section 16. These equations imply that $\nabla^2_{px}\pi(p^*,x^*)x^* = \nabla_p\pi(p^*,x^*)$ and $p^{*T}\nabla^2_{px}\pi(p^*,x^*) = \nabla_x\pi(p^*,x^*)^T$. The same equations will apply to the corresponding partial derivatives of $\pi^*(p^*,x^*)$. Thus we need only satisfy equations (294) for the (M−1) by (N−1) submatrix of the N M matrix $\nabla^2_{px}\pi(p^*,x^*)$ that drops the last row and column of this matrix. Thus for $\pi(p,x)$ to be flexible at (p^*,x^*), we need to satisfy M + N − 1 of the equations in (290) and (291), the M(M−1)/2 equations in the upper triangle of equations (292), the N(N−1)/2 in the upper triangle of equations (293) and the (M−1)(N−1) equations in (294) that drop the equations for one row and one column of the matrix equation involving M rows and N columns. Thus a flexible functional form for a regular variable profit function must have at least M + N − 1 + M(M−1)/2 + N(N−1)/2 + (M−1)(N−1) independent parameters.

Recall that the normalized quadratic unit cost function was defined by (146)-(149) in Section 10 above. We will adapt this functional form to our present context. Define the function r(p) for $p > 0_M$ as follows:

(296) $r(p) \equiv b^Tp + (1/2)p^TBp/\beta^Tp$

where $\beta > 0_M$ is a predetermined vector, b is a parameter vector and B is symmetric positive semidefinite parameter matrix that satisfies:

(297) $Bp^* = 0_M$.

Use the normalized quadratic functional form to define the following function of f(x) for $x > 0_N$:

(298) $f(x) \equiv a^Tx + (1/2)x^TAx/\alpha^Tx$

where $\alpha > 0_N$ is a predetermined vector, a is a parameter vector and A is symmetric negative semidefinite parameter matrix that satisfies:

(299) $Ax^* = 0_N$.

Normalize α and β so that they satisfy the following restrictions:

(300) $\alpha^Tx^* = 1$; $\beta^Tp^* = 1$.

Use the f(x) and r(p) defined above in the following definition for the *normalized quadratic variable profit function*:[109]

(301) $\pi(p,x) \equiv r(p)f(x) + p^TCx$

where C is an M by N parameter matrix. Using the restrictions defined by (297), (299) and (300), the level and first and second order partial derivatives of the $\pi(p,x)$ defined by (300) evaluated at (p^*,x^*) are set equal to the corresponding level and derivatives of an exogenously given $\pi^*(p^*,x^*)$

(302) $\pi^*(p^*,x^*) = a^Tx^*b^Tp^* + p^{*T}Cx^*$;
(303) $\nabla_p\pi^*(p^*,x^*) = ba^Tx^* + Cx^*$;
(304) $\nabla_x\pi^*(p^*,x^*) = ab^Tp^* + C^Tp^*$;
(305) $\nabla^2_{pp}\pi^*(p^*,x^*) = Ba^Tx^*$;
(306) $\nabla^2_{xx}\pi^*(p^*,x^*) = Ab^Tp^*$;
(307) $\nabla^2_{px}\pi^*(p^*,x^*) = ba^T + C$.

We show that there is an a, b, A, B and C solution to the above equations. Tentatively assume that:

(308) $a^Tx^* = 1$; $Cx^* = 0_M$ and $p^{*T}C = 0_N^T$.

Substitute (308) into (303) and solve for $b = \nabla_p\pi^*(p^*,x^*)$. This implies that $p^{*T}b = \pi^*(p^*,x^*)$. Substitute (308) into (304) and solve for $a = \nabla_x\pi^*(p^*,x^*)/b^Tp^* = \nabla_x\pi^*(p^*,x^*)/\pi^*(p^*,x^*)$. Since $x^{*T}\nabla_x\pi^*(p^*,x^*) = \pi^*(p^*,x^*)$, it can be seen that $a^Tx^* = 1$. Substitute this equation into (305) and solve for $B = \nabla^2_{pp}\pi^*(p^*,x^*)$, a symmetric positive semidefinite matrix that satisfies $Bp^* = 0_M$ using the linear homogeneity of $\pi^*(p,x)$ in p. Using $p^{*T}b = \pi^*(p^*,x^*)$, (306) implies that $A = [\pi^*(p^*,x^*)]^{-1}\nabla^2_{xx}\pi^*(p^*,x^*)$. Thus A is a negative semidefinite matrix that satisfies $Ax^* = 0_N$. Finally, define $C \equiv \nabla^2_{px}\pi^*(p^*,x^*) - ba^T = \nabla^2_{px}\pi^*(p^*,x^*) - [\pi^*(p^*,x^*)]^{-1}\nabla_p\pi^*(p^*,x^*)\nabla_x\pi^*(p^*,x^*)^T$. Using $\nabla^2_{px}\pi^*(p^*,x^*)x^* = \nabla_p\pi^*(p^*,x^*)$, $p^{*T}\nabla_p\pi^*(p^*,x^*) = \nabla_x\pi^*(p^*,x^*)^T$ and $x^{*T}\nabla_x\pi^*(p^*,x^*) = \pi^*(p^*,x^*) = p^{*T}\nabla_p\pi^*(p^*,x^*)$, it can be seen that $Cx^* = 0_M$ and $p^{*T}C = 0_N^T$. Thus the normalized quadratic profit function defined by (301) is a flexible functional form.

Given data on net outputs y^t, "fixed" inputs x^t and their prices p^t and w^t for t = 1,...,T, econometric estimating equations for a production unit whose technology is (approximately) dual to the profit function $\pi(p,x)$ defined by (301) can be obtained by using Hotelling's Lemma and Samuelson's Lemma to generate the following nonlinear estimating equations for t = 1,...,T:

(309) $y^t = [b + (\beta^Tp^t)^{-1}Bp^t - (½)(\beta^Tp^t)^{-2}p^{tT}Bp^t\beta][a^Tx + (½)x^{tT}Ax^t/\alpha^Tx^t] + Cx^t + u^t$;

[109] An alternative functional form for a variable profit function that used the r(p) and f(x) defined by (296) and (298) as building blocks appeared in Diewert and Fox (2017). Note that net outputs y and fixed inputs x are *separable* if $C = O_{M\times N}$, an M by N matrix of 0's. See Blackorby, Primont and Russell (1978) on separability concepts.

$$(310)\ w^t = [a + (\alpha^T x^t)^{-1}Ax^t - (½)(\alpha^T x^t)^{-2}x^{tT}Ax^t\alpha][b^Tp + (½)p^{tT}Bp^t/\beta^T p^t] + C^Tp^t + v^t$$

where the error vectors u^t and v^t have zero means. The disadvantage of these estimating equations is that they are more complicated than the rather straightforward comparable translog estimating equations that were obtained in the previous section. However, this functional form has the advantage that the appropriate curvature conditions can be imposed; i.e., the matrices A and B that appear in the above equations can be replaced by $A = - A^*A^{*T}$ and $B = B^*B^{*T}$ where A^* and B^* are lower triangular matrices with $A^{*T}x^* = 0_N$ and $B^{*T}p^* = 0_M$.[110] These substitutions will not destroy the flexibility of the resulting functional form. Semiflexible versions of the A and B matrices can also be estimated in order to conserve on the number of parameters in the model. Finally, technical progress can easily be accommodated in the above model: simply add the time trend vector a^*t to the a vector and add the time trend vector b^*t to the vector b in the estimating equations (308) and (309) for period t.[111]

19. The KBF Variable Profit Function

In Section 11 of this chapter, we studied the KBF unit cost function. This functional form can be used as a basic building block to obtain a flexible functional form for a variable profit function that is dual to a regular production possibilities set.[112] Thus define the function r(p) for $p > 0_M$ as follows:

$$(311)\ r(p) \equiv [p^T(bb^T + B)p]^{1/2}$$

where b is a parameter vector and B is symmetric positive semidefinite parameter matrix that satisfies:

$$(312)\ Bp^* = 0_M.$$

We also use the KBF functional form to define the following function of f(x) for $x > 0_N$:

$$(313)\ f(x) \equiv [x^T(aa^T + A)x]^{1/2}$$

where a is a parameter vector and A is symmetric negative semidefinite parameter matrix that satisfies:

$$(314)\ Ax^* = 0_N.$$

[110] After making these substitutions for A and B, the resulting $\pi(p,x)$ will satisfy the convexity and concavity conditions at the point (p,x) provided that $p > 0_M$, $x > 0_N$, $r(p) > 0$ and $f(x) > 0$.

[111] For identification, add the constraint $a^{*T}1_N = 0$. Of course, to achieve additional flexibility, linear or quadratic splines in time could be added to the a and b vectors; see Fox (1998) or Fox and Grafton (2000) for empirical examples using the normalized quadratic functional form and piece-wise linear splines to model technical progress.

[112] The advantage of using this functional form over using the normalized quadratic as a basic building block is that when using the KBF functional form, we do not have to specify the exogenous vectors α and β which appeared in the normalized quadratic functional form.

We use the f(x) and r(p) defined above in the following definition for the *KBF variable profit function*:[113]

(315) $\pi(p,x) \equiv r(p)f(x) + p^TCx$

where C is an M by N parameter matrix. Using the restrictions defined by (312) and (314), the level and first and second order partial derivatives of the $\pi(p,x)$ defined by (315) evaluated at (p^*,x^*) are set equal to the corresponding level and derivatives of an exogenously given $\pi^*(p^*,x^*)$

(316) $\pi^*(p^*,x^*) = a^Tx^*b^Tp^* + p^{*T}Cx^*$;
(317) $\nabla_p\pi^*(p^*,x^*) = ba^Tx^* + Cx^*$;
(318) $\nabla_x\pi^*(p^*,x^*) = ab^Tp^* + C^Tp^*$;
(319) $\nabla^2_{pp}\pi^*(p^*,x^*) = Ba^Tx^*/b^Tp^*$;
(320) $\nabla^2_{xx}\pi^*(p^*,x^*) = Ab^Tp^*/a^Tx^*$;
(321) $\nabla^2_{px}\pi^*(p^*,x^*) = ba^T + C$.

It can be seen that these equations are identical to equations (302)-(307) in the previous section except that equations (319) and (320) are slightly different from the corresponding equations (305) and (306). It turns out that this difference does not affect the proof that there is an a, b, A, B and C solution to the above equations. Thus it is straightforward to establish that the KBF variable profit function is a flexible functional form.

Given data on net outputs y^t, "fixed" inputs x^t and their prices p^t and w^t for t = 1,...,T, econometric estimating equations for a production unit whose technology is (approximately) dual to the profit function $\pi(p,x)$ defined by (315) can be obtained by using Hotelling's Lemma and Samuelson's Lemma to generate the following nonlinear estimating equations for t = 1,...,T:

(322) $y^t = [bb^Tp^t + Bp^t][p^{tT}(bb^T + B)p^t]^{-1/2}[x^{tT}(aa^T + A)x^t]^{1/2} + Cx^t + u^t$;
(323) $w^t = [aa^Tx^t + Ax^t][x^{tT}(aa^T + A)x^t]^{-1/2}[p^{tT}(bb^T + B)p^t]^{1/2} + C^Tp^t + v^t$

where the error vectors u^t and v^t have zero means. Again, the disadvantage of these estimating equations is that they are a lot more complicated than the rather straightforward comparable translog estimating equations that were obtained for the translog functional form. However, as was the case with the normalized quadratic profit function, this functional form has the advantage that the appropriate curvature conditions can be imposed without destroying the flexibility of the functional form; i.e., the matrices A and B that appear in the above equations can be replaced by $A = -A^*A^{*T}$ and $B = B^*B^{*T}$ where A^* and B^* are lower triangular matrices with $A^{*T}x^* = 0_N$ and $B^{*T}p^* = 0_M$.[114] As usual, semiflexible versions of the A and B matrices can also be estimated in order to

[113] Net outputs y will be separable from inputs x if $C = O_{M\times N}$.

[114] After making these substitutions for A and B, the resulting $\pi(p,x)$ will satisfy the convexity and concavity conditions provided at the point (p,x) provided that $p > 0_M$, $x > 0_N$, r(p) > 0 and f(x) > 0.

conserve on the number of parameters in the model. And again as usual, flexible forms of technical progress can easily be accommodated in the above model by adding the time trend vector a^*t to the a vector and add the time trend vector b^*t to the vector b in the estimating equations (322) and (323) for period t.[115]

The KBF functional form developed in this section is very similar to the normalized quadratic functional form that was developed in the previous section. However, the KBF functional form has the advantage that it is not necessary to specify an α and β vector a priori as was the case for the normalized quadratic profit function. The KBF functional form seems to be the most promising parsimonious functional form that has been developed up to the present.

20. Joint Cost Functions

Instead of maximizing profits with respect to variable inputs and outputs, in this section we minimize cost subject to producing a specified vector of outputs. Thus consider a production unit that produces the output vector $y \geq 0_M$ using an input vector $x \geq 0_N$. The set of feasible output and input vectors (y,x) is a set S which satisfies the following minimal regularity condition:[116]

(324) S is closed subset of M + N space such that for every output vector $y \geq 0_M$, there exists an input vector $x \geq 0_N$ such that $(y,x)\in S$.

Let $w >> 0_N$ be a strictly positive vector of input prices and let $y \geq 0_M$ be an output vector. Define the producer's *joint cost function* C(y,w) as follows:

(325) $C(y,w) \equiv \min_x \{w^Tx : (y,x)\in S\}$.

The regularity conditions (324) on S and the assumption that $w >> 0_N$ imply that the minimum in (325) will exist.

It is frequently useful to assume that S satisfies *free disposability of inputs*, property (326) below, and/or *free disposability of outputs*, property (327) below.

(326) $y \geq 0_M$, $0_N \leq x^1 < x^2$ and $(y,x^1)\in S$ implies $(y,x^2)\in S$.
(327) $0_M \leq y^1 < y^2$ and $(y^2,x)\in S$ implies $(y^1,x)\in S$.

Problems

27. *Theorem 15*: Suppose S satisfies conditions (324) and define C(y,w) by (325) for $y \geq 0_M$ and $w >> 0_N$. Show that C(y,w) has the following properties:

---

[115] Again, in order to identify all of the parameters, add the constraint $a^{*T}1_N = 0$. To achieve additional flexibility, linear or quadratic splines in time could be added to the a and b vectors.
[116] Note that in this section, y is a vector of outputs rather than a vector of net outputs as in previous sections.

(i) $C(y,w)$ is a *nonnegative* function; i.e., $C(y,w) \geq 0$ for $y \geq 0_M$ and $w >> 0_N$.

(ii) $C(y,w)$ is *positively linearly homogeneous in p* for each fixed y; i.e., $C(y,\lambda w) = \lambda C(y,w)$ for all $\lambda > 0$, $w >> 0_N$ and $y \geq 0_M$.

(iii) $C(y,w)$ is *nondecreasing in w* for each fixed y; i.e., $C(y,w^1) \leq C(y,w^2)$ for $y \geq 0_M$ and $w^2 > w^1 >> 0_N$.

(iv) $C(y,w)$ is a *concave function of w* for each fixed y; i.e., $C(y,\lambda w^1+(1-\lambda)w^2) \geq \lambda C(y,w^1) + (1-\lambda)C(y,w^2)$ for $y \geq 0_M$, $w^1 >> 0_N$; $w^2 >> 0_N$ and $0 < \lambda < 1$.

(v) $C(y,w)$ is a *continuous function of w* for each fixed $y \geq 0_M$.

Hint: Adapt the proof of Theorem 1 in Section 2 above.

28. Continuation of 27: Suppose S satisfies the free disposability of outputs property (327) in addition to the minimal regularity conditions (324). Show that $C(y,w)$ is *nondecreasing in y* for fixed w; i.e., show that $w >> 0_N$, $0_M \leq y^1 < y^2$ and $(y^2,x)\in S$ implies $C(y^1,w) \leq C(y^2,w)$. *Hint*: Use a feasibility argument.

Thus the joint cost function $C(y,w)$ has much the same properties with respect to input prices as the single output cost function that was studied in Section 2 above. In particular, $C(y,w)$ *must* be linearly homogeneous and concave in w for fixed y.

Under what conditions can a knowledge of the joint cost function, $C(y,w)$, be sufficient to determine the underlying technology set S? We now address this question. Suppose S satisfies the minimal regularity conditions (324). For each $y \geq 0_M$, define the set of inputs that can produce at least y, $L(y)$, as follows:

(328) $L(y) \equiv \{x : (y,x)\in S\}$.

If we are given the family of upper level sets, $L(y)$ for every $y \geq 0_M$, then S can be recovered using $S = \{(y,x): y \geq 0_M \text{ and } x\in L(y)\}$. Thus the above question can be reduced to the equivalent question: under what assumptions on $L(y)$ can the joint cost function be used to determine $L(y)$ for each $y \geq 0_M$? We can use the method explained in Section 3 above to answer this question.

Let $y \geq 0_M$ and $w >> 0_N$. Use the given joint cost function $C(y,w)$ to define the following half space of inputs:

(329) $M(y,w) \equiv \{x: w^Tx \geq C(y,w)\}$.

The above half space must contain the level set $L(y)$. Thus $L(y)$ must be contained in the following set, which is the intersection of all of the supporting halfspaces to $L(y)$:

(330) $M(y) \equiv \bigcap_{w>>0_N} M(y,w)$.

Since each of the sets in the intersection, M(y,w), is a convex set, then M(y) is also a convex set. Since L(y) is a subset of each M(y,w), it must be the case that L(y) is also a subset of M(y); i.e., we have $L(y) \subset M(y)$. As was the case in Section 3, in order to ensure that M(y) = L(y), we need to add the following two conditions on the family of level sets L(y):

(331) For each $y \geq 0_M$, L(y) satisfies free disposability of inputs; i.e., $x^1 \in L(y)$, $x^2 \geq x^1$ implies $x^2 \in L(y)$.

(332) For each $y \geq 0_M$, L(y) is a convex set.

Condition (331) on the family of as input level sets L(y) is equivalent to condition (326) on the production possibilities set S. As in Section 3, assumptions (331) and (332) rule out backward bending and nonconvex input production possibilities sets L(y).

As was the case in Section 3, if the producer is a price taker in input markets, then it is not necessary to assume properties (331) and (332) when estimating a joint cost function: a cost minimizing producer will never choose an input vector that belongs to a nonconvex or backward bending upper level set L(y). Thus an estimated joint cost function can be used to form the upper level sets M(y) and these sets can provide an adequate approximation to the true L(y) for most purposes.

If the joint cost function C(y,w) satisfies the conditions listed in Theorem 15 and is differentiable with respect to input prices w, then we can show that Shephard's Lemma still holds; i.e., the producer's system of cost minimizing input demand functions is equal to $x(y,w) \equiv \nabla_w C(y,w)$ for $y \geq 0_M$ and $w >> 0_N$.[117]

If the production possibilities set S has additional properties, then we can deduce that the joint cost function C(y,w) has additional properties. Two familiar additional properties for S are the following ones:

(333) S is a convex set; i.e., $(y^1,x^1) \in S$, $(y^1,x^1) \in S$ and $0 < \lambda < 1$ implies $(\lambda y^1 + (1-\lambda)y^2, \lambda x^1 + (1-\lambda)x^2) \in S$.

(334) S is a cone; i.e., if $(y,x) \in S$ and $\lambda > 0$, then $(\lambda y, \lambda x) \in S$.

The cone assumption (334) means that production is subject to constant returns to scale. The convexity assumption rules out technologies that are subject to increasing returns to scale. Some of the implications of these assumptions are listed in the following problems.

Problems

[117] The proof of Theorem 5 in Section 4 can be adapted to prove this result.

29. Assume S satisfies (234) and the convexity assumption (333). (i) Show that $L(y) \equiv \{x : (y,x)\in S\}$ is a convex set for each $y \geq 0_M$. (ii) Show that C(y,w) defined by (325) is a convex function of y for fixed $w >> 0_N$. *Hint*: Look at the proof of part (b) of Theorem 11.

30. Assume S satisfies (234) and the output free disposability assumption (327). Show that C(y,w) defined by (325) is a nondecreasing function of y for fixed $w >> 0_N$. *Hint*: Use a feasibility argument.

31. Assume S satisfies (234) and the cone assumption (334). Show that C(y,w) defined by (325) is a linearly homogeneous function of y for fixed $w >> 0_N$. *Hint*: Modify the proof of part (c) of Theorem 11.

The above problems show that if S satisfies the output free disposal assumption (327) and the convexity and constant returns to scale assumptions (333) and (334), then the corresponding joint cost function C(y,w) will be a nondecreasing, linearly homogeneous and convex function of y for fixed w.

Assume that C(y,w) is differentiable with respect to y and w. Shephard's Lemma enables us to interpret the vector of first order partial derivatives of the joint cost function with respect to the input price vector w, $\nabla_w C(y,w)$, as the producer's vector of input demand functions, x(y,w). The vector of first order partial derivatives of the joint cost function with respect to y, $\nabla_y C(y,w)$, is obviously the vector of *marginal costs* for each output. However, if S satisfies the convexity assumption (333), then $p = \nabla_y C(y,w)$ can be interpreted as the producer's *system of inverse supply functions*; i.e., if the producer faced the output price vector p and the input price vector w, then an output vector y which satisfied the system of equations $p = \nabla_y C(y,w)$ and the $x = \nabla_w C(y,w)$ would be a solution to the following producer's profit maximization problem:

(335) $\max_{y,x} \{p^T y - w^T x : (y,x)\in S\}$.

Theorem 16: Suppose the technology set S satisfies the minimal regularity assumptions (324) plus (326) (free disposability of inputs), (327) (free disposability of outputs) and (333) (convexity). Let $y^* \geq 0_M$ and $w^* >> 0_N$. Suppose that C(y,w) is differentiable at (y^*,w^*). Define $x^* \equiv \nabla_w C(y^*,w^*)$ and $p^* \equiv \nabla_y C(y^*,w^*)$. Then (y^*,x^*) is a solution to the following profit maximization problem:

(336) $\max_{y,x} \{p^{*T} y - w^{*T} x : (y,x)\in S\}$.

Proof: The free disposability assumptions imply that $p^* \geq 0_M$ and $x^* \geq 0_N$. The convexity assumption on S implies that $C(y,w^*)$ is a convex function of y. Thus the function $f(y) \equiv C(y,w^*) - p^{*T} y$ is also a convex function of y for all $y \geq 0_M$. Note that $\nabla f(y^*) = \nabla_y C(y^*,w^*) - p^* = 0_M$ using the definition of p^*. Since f(y) is a convex function and differentiable at $y = y^*$, its first order Taylor series approximation around this point will lie below (or be coincident with) f(y). Thus we have for all $y \geq 0_M$:

$$(337)\ f(y) \geq f(y^*) + \nabla f(y^*)^T(y-y^*) = f(y^*)$$

where the inequality follows since $\nabla f(y^*) = 0_M$. Thus f(y) attains a global minimum at y^*. Using the definition of f, we see that y^* is a solution to the following minimization problem:

$$(338)\ \min_y \{C(y,w^*) - p^{*T}y\ ;\ y \geq 0_M\} = C(y^*,w^*) - p^{*T}y^* = w^{*T}\nabla_w C(y^*,w^*) - p^{*T}y^* = w^{*T}x^* - p^{*T}y^*$$

where the second equality follows from the linear homogeneity of C(y,w) in w and the third equality follows from the definition of $x^* \equiv \nabla_w C(y^*,w^*)$.

It can be verified that solving the profit maximization problem defined by (335) is equivalent to solving the following *(net) cost minimization problem*:

$$(339)\ \min_{y,x} \{w^{*T}x - p^{*T}y : (y,x)\in S\} = \min_y \{[\min_x w^{*T}x : (y,x)\in S] - p^{*T}y\}$$
$$= \min_y \{C(y,w^*) - p^{*T}y\ ;\ y \geq 0_M\} \qquad \text{using (325)}$$
$$= w^{*T}x^* - p^{*T}y^* \qquad \text{using (338).}$$

Q.E.D.

The above result is a joint cost function counterpart to Samuelson's Lemma, Theorem 12 above. It says that if producers take prices as given on both input and output markets and the technology set is convex, then the producer's system of inverse supply functions, p(y,x) is equal to $\nabla_y C(y,w)$, the producer's system of marginal cost functions.

If the production possibilities set S satisfies all of the regularity conditions on S that are listed in this section (free disposability of inputs and outputs, convexity and constant returns to scale), we say that S is a *regular production possibilities set*.

Problems

32. Suppose S satisfies the minimal regularity conditions (324). Define the corresponding joint cost function C(y,w) by (325). Suppose C(y,w) is twice continuously differentiable with respect to w at some point $y > 0_M$ and $w >> 0_N$. Then the system of cost minimizing input demand functions is $x(y,w) = \nabla_w C(y,w)$ and and the N by N matrix of demand derivatives with respect to input prices, $B \equiv [\partial x_n(y,w)/\partial w_i] = \nabla^2_{ww} C(y,w)$ exists. Show that the matrix B has the following properties:

(i) $B = B^T$;
(ii) B is negative semidefinite and
(iii) $Bw = 0_M$.

Hint: Adapt the proof of Theorem 13.

33. Suppose S is a regular production possibilities set and the corresponding C(y,w) is twice continuously differentiable at the point $y >> 0_M$ and $w >> 0_N$. Then the system of inverse supply functions, $p(y,w) = \nabla_y C(y,w)$ and the M by M matrix of partial derivatives with respect to output quantities, $A \equiv [\partial p_m(y,w)/\partial y_k] = \nabla^2_{yy}C(y,w)$ exist. Show that the matrix A has the following properties:

(i) $A = A^T$ so that $\partial p_m(y,w)/\partial y_k = \partial p_k(y,w)/\partial y_m$ for all $m \neq k$;
(ii) A is positive semidefinite and
(iii) $Ay = 0_M$.

Hint: Adapt the proof of Theorem 14.

34. Suppose S is a regular production possibilities set and the corresponding C(y,w) is twice continuously differentiable at the point $y >> 0_M$ and $w >> 0_N$. Then the system of inverse supply functions is $p(y,w) = \nabla_y C(y,w)$ and the M by N matrix of partial derivatives supply prices with respect to input prices, $D \equiv [\partial p_m(y,w)/\partial w_n] = \nabla^2_{yw}C(y,w)$ exists. The system of cost minimizing input demand functions is $x(y,w) = \nabla_w C(y,w)$ and the N by M matrix of partial derivatives of input quantities with respect to output quantities, $E \equiv [\partial x_n(y,w)/\partial y_m] = \nabla^2_{wy}C(y,w)$ exists. Show that the matrices D and E have the following properties:

(i) $D = E^T$;
(ii) $p(y,w) = Dw \geq 0_M$;
(ii) $x(y,w) = Ey \geq 0_N$.

Hint: Adapt the proof of Theorem 15.

Shephard's Lemma and Theorem 16 can be used as a convenient method for obtaining econometric estimating equations for determining the parameters that characterize a producer's technology set S. Assuming that S satisfies the minimal regularity conditions on S, we need only postulate a differentiable functional form for the producer's joint cost function, C(y,w), that is linearly homogeneous and concave in w. Suppose that we have collected data on the input vectors used by the unit in period t, x^t, and the outputs produced in period t, y^t, for t = 1,…,T time periods as well as the corresponding input price vectors w^t. Then the following NT equations can be used in order to estimate the unknown parameters in C(y,w):

(340) $x^t = \nabla_w C(y^t,w^t) + u^t$; $\qquad t = 1,\ldots,T$

where u^t is a vector of errors. If in addition, S is a convex set and the firm is maximizing profits facing the fixed output and input price vectors, p^t and w^t, respectively in period t, then the following MT equations can be added to (340) as additional estimating equations:

(341) $p^t = \nabla_y C(y^t,w^t) + v^t$; $\qquad t = 1,\ldots,T$

where v^t is a vector of errors.[118]

21. Flexible Functional Forms for Joint Cost Functions

Specific functional forms for C(y,w) can be found by adapting the functional forms explained in Sections 17-19 above. Adapting the material in Section 17, we could assume the log of the joint cost function for a regular technology, lnC(y,w), has the following *translog* functional form:[119]

$$(342)\ \ln C(y,w) \equiv a_0 + \sum_{m=1}^{M} a_m \ln y_m + (1/2) \sum_{m=1}^{M}\sum_{k=1}^{M} a_{mk} \ln y_m \ln y_k + \sum_{n=1}^{N} b_n \ln w_n + (1/2)\sum_{n=1}^{N}\sum_{j=1}^{N} b_{ni} w_n \ln w_i + \sum_{m=1}^{M}\sum_{n=1}^{N} c_{mn} \ln y_m \ln w_n\ .$$

The unknown coefficients in (342) must satisfy the restrictions (276)-(283) listed in Section 17 if S is a regular production possibilities set.

Note that using Shephard's Lemma, we have $\partial \ln C(y,w)/\ln\partial w_n = [w_n/C(y,w)]\partial C(y,w)/\partial w_n = [w_n/C(y,w)]x_n(y,w) \equiv S_n(y,w)$ where $x_n(y,w)$ is the cost minimizing demand function for input and $S_n(y,w)$ is the *share of input* n *in total cost*. Assuming that the producer minimizes cost and S is dual to the translog joint cost function defined by (342), then differentiating the logarithm of C(y,w) defined by (342) with respect to the logarithm of w_n leads to the following system of *input share equations*:

$$(343)\ S_n(y,w) = b_n + \sum_{j=1}^{N} b_{nj} \ln w_j + \sum_{m=1}^{M} c_{mn} \ln y_m\ ; \qquad n = 1,...,N.$$

Equations (342) and (343) can be used as estimating equations if the production unit is minimizing costs. Note that these equations are linear in the unknown parameters.[120]

Suppose that in addition to the assumption that the production unit is minimizing costs, we assume that the technology set is regular and the producer is maximizing profits. Using Theorem 16, $\partial \ln C(y,w)/\ln y_m = [y_m/C(y,w)]\partial C(y,w)/\partial y_m = [y_m/C(y,w)]p_m(y,w) \equiv s_m(p,x)$ where $p_m(y,w)$ is the profit maximizing inverse demand function for output m and $s_m(p,x)$ is the share of output m in total profit maximizing revenue. Assume that S is regular. Assuming that the producer maximizes profit and S is dual to the translog joint cost function C(y,w) defined by (342), then differentiating the logarithm of C(y,w) with respect to the logarithm of y_m leads to the following system of *revenue share equations*:

$$(344)\ s_m(y,w) = a_m + \sum_{k=1}^{M} a_{mk} \ln y_k + \sum_{n=1}^{N} c_{mn} \ln w_n\ ; \qquad m = 1,...,M.$$

---

[118] If in addition, the technology set S is subject to constant returns to scale and the data reflect this fact by satisfying $p^{tT}y^t = w^{tT}x^t$ for t = 1,…,T, then the error vectors u^t and v^t in (340) and (341) cannot be statistically independent. Hence one of the M+N equations in (340) and (341) must be dropped from the system of estimating equations.

[119] This functional form is due to Burgess (1974) who applied it to international trade theory. For applications of this functional form to index number theory, see Diewert and Morrison (1986) and Diewert and Fox (2010).

[120] If we do not impose constant returns to scale and convexity on S, then the parameter restrictions (277) and (281)-(283) do not have to be imposed. These restrictions should be imposed if we assume constant returns to scale and convexity.

Equations (342)-(344) can be used as estimating equations if the production unit is maximizing profits and has a regular translog technology.

The above functional form for the logarithm of joint cost does not allow for technical progress. To remedy this problem, simply add the following terms to the right hand side of definition (342): $\alpha_0 t + \sum_{m=1}^{M} t\alpha_m \ln y_m + \sum_{n=1}^{N} t\beta_n \ln w_n$ where t is a scalar time variable and the new parameters α_m and β_n satisfy $\sum_{m=1}^{M} \alpha_m = 0$ and $\sum_{n=1}^{N} \beta_n = 0$.[121]

A problem with the translog joint cost function is that it is not possible to impose concavity in w (and convexity in y if the dual S satisfies convexity) over the region spanned by the sample input prices w^t (and the region spanned by the sample output vectors y^t if S is a convex set) without impairing the flexibility of the functional form. In order to impose these curvature conditions without destroying the flexibility property, we turn to the functional forms defined in Sections 18 and 19.

Define the *normalized quadratic joint cost function* $C(y,w)$ for $y > 0_M$ and $w > 0_N$ as follows:

$$(345)\quad C(y,w) \equiv g(y)c(w) + y^T E w$$

where $g(y) \equiv b^T y + (1/2)y^T By/\beta^T y$, $\beta > 0_M$ is a predetermined vector that satisfies $\beta^T y^* = 1$, $b > 0_M$ is a parameter vector, B is symmetric positive semidefinite parameter matrix that satisfies $By^* = 0_M$, $c(w) \equiv a^T w + (1/2)w^T Aw/\alpha^T w$, $\alpha > 0_N$ is a predetermined vector that satisfies $\alpha^T w^* = 1$, a is a parameter vector that satisfies $a^T w^* = 1$, A is symmetric negative semidefinite parameter matrix that satisfies $Aw^* = 0_N$ and E is an M by N parameter matrix.

Define the *KBF joint cost function* $C(y,w)$ using (345) where E is again an M by N parameter matrix. However, redefine g(y) and c(w) as follows: $g(y) \equiv (y^T [bb^T + B]y)^{1/2}$, where $b > 0_M$ is a parameter vector, B is symmetric positive semidefinite parameter matrix that satisfies $By^* = 0_M$, $c(w) \equiv (w^T[aa^T + A]w)^{1/2}$, a is a parameter vector that satisfies $a^T w^* = 1$ and A is symmetric negative semidefinite parameter matrix that satisfies $Aw^* = 0_N$.

For both of these joint cost functions, the vector of cost minimizing input demand functions x(y,w) can be obtained by calculating the vector of first order partial derivatives, $\nabla_w C(y,w)$. The concavity in input prices property for the joint cost function can be imposed by setting $A = - A^* A^{*T}$ with A^* lower triangular and $A^{*T} w^* = 0_N$. In the case where the underlying production possibilities set S is convex, the vector of profit maximizing output prices p(y,w) that is consistent with the production of the vector y of outputs can be obtained by calculating the vector of first order partial derivatives,

[121] Linear or quadratic spline functions in time can also be added to the estimating equations to better approximate variable rates of technical progress over time.

$\nabla_y C(y,w)$. The convexity property in output quantities for C(y,w) can be imposed by setting $B = B^* B^{*T}$ with B^* lower triangular and $B^{*T} y^* = 0_M$.[122]

The normalized quadratic and KBF joint cost functions as defined above, do not allow for technical progress. This problem can be remedied by adding the term $(\Sigma_{m=1}^M \gamma_m y_m t)(\Sigma_{n=1}^N \delta_n w_n t)$ to the right hand side of definitions (345) where the γ_m and δ_n are technical progress parameters and t is a time trend.[123] These additional technical progress terms may not capture the trends in technical progress in the time series context if the sample period is long. In this case, the terms $\Sigma_{m=1}^M \gamma_m y_m t$ and $\Sigma_{n=1}^N \delta_n w_n t$ can be replaced by piece-wise linear spline functions as was done in Section 13 above; see equations (200).

22. Applications of Joint Cost Functions

In this section, we discuss three areas of research where joint cost functions play important roles.

Many government outputs are produced in a nonmarket context. The output quantities can usually be measured but typically, there are no market prices for the outputs that are produced by many government production units. However, government producers still have an incentive to minimize costs. If the public sector production unit is minimizing costs and the technology set can be approximated by a constant returns to scale production possibilities set S and econometric estimation of a differentiable dual joint cost function C(y,w) is possible (using just the input demand functions as estimating equations), then approximate output prices can be obtained as the vector of marginal costs, $p \equiv \nabla_y C(y,w)$. If production is subject to constant returns to scale, then the resulting output price vector p will have the property that $p^T y = C(y,w) = w^T x$; i.e., the resulting value of outputs will equal the value of inputs.[124] This result is useful in the national income accounting context where government statisticians have to find methods for valuing public sector outputs. Using marginal cost prices is also useful when economists want to measure the productivity performance of public sector production units.[125]

A second application for the estimation of joint cost functions is in the context of the regulation of utilities that deliver electricity, water and communications services via networks. Regulators are interested in using marginal costs to aid them in setting utility prices. Utilities may be forced to sell their outputs at regulated prices that do not reflect marginal costs but regulated utility firms will still have an incentive to minimize costs. In

---

122 After making these substitutions for A and B, the resulting C(y,w) will satisfy the convexity and concavity conditions provided at the point (y,w) provided that $y > 0_M$, $w > 0_N$, $g(y) > 0$ and $c(w) > 0$. The proof of the flexibility of the normalized quadratic and KBF joint cost functions in the case of a regular technology is entirely analogous to the corresponding proofs of normalized quadratic and KBF variable profit functions that were discussed in Sections 18 and 19.

123 In order to identify all of these technical progress parameters, we need to impose a normalization on them such as $\Sigma_{m=1}^M \gamma_m = 1$.

124 In practice, the vector of marginal costs may have to be approximated by average costs of production, which in turn will usually require many accounting imputations.

125 See Diewert (2011) (2012) (2018) on this topic.

this case, joint cost functions can be estimated and the resulting estimates can be used to measure technical progress as well as the Total Factor Productivity of the regulated firms.[126]

A third area where joint cost functions play an important role is in modeling monopolistic behavior. Typically producers take input prices as fixed and beyond their control. However, they may have some pricing power over their outputs. Recall (245) which defined a producer's competitive profit maximization problem. A *monopolistic counterpart* to this problem is the following problem:

$$(346)\ \max_{y,x} \{\Sigma_{m=1}^{M} f_m(y_m)y_m - w^Tx : (y,x)\in S\} = \max_{y} \{\Sigma_{m=1}^{M} f_m(y_m)y_m - C(y,w)\}$$

where $w >> 0_N$ is a positive input price vector, $y \equiv (y_1,...,y_M) \geq 0_M$ is an output vector, S is the producer's production possibilities set, C(y,w) is the producer's joint cost function defined by (325) and $p_m = f_m(y_m)$ is the (downward sloping) *inverse demand function* for output m that the producer faces for m = 1,...,M. If the inverse demand functions $f_m(y_m)$ and the joint cost function C(y,w) are once differentiable when evaluated at the period t data, then under appropriate regularity conditions on the $f_m(y_m)$ and S, the following equations will be satisfied by a profit maximizing monopolist using the observed period t data:

$$(347)\ p^t(1_M - \mu^t) = \nabla_y C(y^t,w^t)\ ; \qquad t = 1,...,T;$$
$$(348)\ x^t = \nabla_w C(y^t,w^t)\ ; \qquad t = 1,...,T$$

where 1_M is an M dimensional vector of ones, $\mu^t \equiv [\mu_1^t,...,\mu_M^t]^T \geq 0_M$ is a period t *markup vector* where $\mu_m^t \equiv -[y_m^t/p_m^t][\partial f_m(y_m^t)/\partial y_m]$ is the markup of price over marginal cost for output m in period t, y^t and x^t are the observed quantity vectors for outputs and inputs in period t and p^t and w^t are the corresponding observed output and input price vectors for period t = 1,...,T. If the markups are constant over time, given a suitable functional form for the joint cost function C(y,w), equations (347) and (348) can be used as econometric estimating equations.[127] Thus again, joint cost functions play a crucial role in this area of economics.[128]

23. Problems that Require Additional Research

We conclude this chapter with some comments on three problem areas that have not been addressed in the above sections.

[126] For examples of the use of joint cost functions in a regulatory context, see Denny, Fuss and Waverman (1981), Lawrence and Diewert (2006) and Diewert, Lawrence and Fallon (2009).

[127] If the markups are not constant, then linear (or piece-wise linear) trends in the markups could be introduced into the model. See Diewert and Fox (2008) for an econometric application of this model and Diewert and Fox (2010) for an application of this model to index number theory.

[128] If the monopolist provides some goods and services on a competitive basis (i.e., at marginal cost), then the markup for this commodity can be set equal to zero. Alternatively, this commodity could be removed from the y vector and be placed with the x inputs, except the quantity would be indexed with a negative sign in the input demand equations. The resulting input cost would become input cost less the revenue from the sales of goods and services provided at marginal cost.

The first problem area is the difficulty of *distinguishing increasing returns to scale from technical progress* if there is general growth of all inputs and outputs for the production unit that is under consideration. Multicollinearity problems usually arise in this situation: the two effects typically cannot be reliably determined using just time series data.

The second problem area is the fact that many inputs cannot be varied in the short run and thus producers are not necessarily producing outputs and utilizing inputs on the frontiers of their production possibilities sets. For example, suppose a recession occurs in the economy so that demand for the outputs of production units declines. Producers can reduce the demand for their variable inputs but they are more or less stuck with their structure inputs and with other durable capital investments that are "bolted down".[129] Thus producers end up being in the interior of their production possibilities sets.[130] When a producer makes an irreversible investment, the total cost of the investment should not be charged to the period when the investment was made but this cost should be allocated over the useful life of the investment. But how exactly should this cost be allocated? This is the *fundamental problem of accounting*.[131] Note that in addition to structure and network capital inputs, a successful R&D project is another example of a fixed cost input whose input cost must be allocated over time in some manner. If there is only a single sunk cost input (or we aggregate all sunk cost inputs into a single input), then it is possible to set up an intertemporal profit maximization problem that justifies the purchase of the fixed input. The price of this fixed asset at a particular point in time is the discounted net revenue generated by the project over its remaining useful life and if this information on discounted net revenues can be forecasted, then the initial cost of the asset can be amortized in a manner that is proportional to the forecasted net revenues by period.[132]

The final problem area that has not been addressed in this survey of the applications of duality theory in production theory is the *new goods problem* and the problem of *quality change*. Modern economies are subject to tremendous product churn, and in addition, revolutionary new products are constantly being developed.[133] Up to this point, we have assumed that the production unit is producing M outputs and N inputs and this set of outputs and inputs remains constant over time (if we are in the time series context) or it remains constant over different production units in the same industry (in the cross

[129] Some labour hoarding may also occur; i.e., the costs of firing and then rehiring workers after the recession is over may be higher than just keeping the workers employed.

[130] This inefficiency problem will be addressed in other chapters in this Handbook using nonparametric production analysis or Data Envelopment Analysis; see Charnes and Cooper (1985). Most of the research in this area is applied to cross sectional or panel data. For an application of the nonparametric approach to production theory and the measurement of efficiency in the time series context, see Diewert and Fox (2018).

[131] See Cairns (2013).

[132] For examples of this methodology, see Diewert (2009), Diewert, Lawrence and Fallon (2009), Diewert and Huang (2011), Cairns (2013) and Diewert and Fox (2016).

[133] See Broda and Weinstein (2006) (2010), Bernard, Redding and Schott (2010; 82) and Hottman, Redding and Weinstein (2016; 1300) for information on the number of products sold in the US (at least 1.6 million). The last three papers have information on the frequency of product entry and exit in the US (about 2% per month).

sectional context). If the underlying technology set S^t for a production unit does not change very much when new outputs appear and some old outputs disappear in period t, then the various econometric models proposed above could in theory deal with this problem if we allow for technical change. But if there are many such changes over many periods, obviously, we will not be able to estimate flexible functional forms due to the proliferation of parameters. Even if output changes are infrequent, the production of a new output and the discontinuance of an existing output could lead to a radical change in the use of inputs as the newer technology replaces the existing one and again, we will have a proliferation of parameters, a lack of degrees of freedom and our suggested econometric approaches will fail. Thus there is a need for further research to address these problems.

References

Allen, R.G.C. (1938), *Mathematical Analysis for Economists*, London: Macmillan.

Arrow, K.J., H.B. Chenery, B.S. Minhas and R.M. Solow (1961), "Capital-Labor Substitution and Economic Efficiency", *Review of Economic Statistics* 63, 225-250.

Berge, C. (1963), *Topological Spaces*, New York: MacMillan.

Bernard, A.B., S.J. Redding and P.K. Schott (2010), "Multiple-Product Firms and Product Switching", *American Economic Review* 100, 70-97.

Blackorby, C. and W.E. Diewert (1979), "Expenditure Functions, Local Duality and Second Order Approximations", *Econometrica* 47, 579-601.

Blackorby, C., D. Primont and R.R. Russell (1978), *Duality, Separability and Functional Structure: Theory and Economic Applications*, New York: North-Holland.

Broda, C. and D.E. Weinstein (2006), "Globalization and the Gains from Variety", *Quarterly Journal of Economics* 121, 541–586.

Broda, C. and D.E. Weinstein (2010), "Product Creation and Destruction: Evidence and Price Implications", *American Economic Review* 100:3, 691-723.

Burgess, D. (1974), "A Cost Minimization Approach to Import Demand Equations", *Review of Economics and Statistics* 56, 224–234.

Cairns, R.D. (2013), "The Fundamental Problem of Accounting", *Canadian Journal of Economics* 46, 634-655.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982a), "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity", *Econometrica* 50, 1393-1414.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982b), "Multilateral Comparisons of Output, Input and Productivity using Superlative Index Numbers", *Economic Journal* 96, 659-679.

Charnes, A. and W.W. Cooper (1985), "Preface to Topics in Data Envelopment Analysis", *Annals of Operations Research* 2, 59-94.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1971), "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica* 39, 255-256.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau (1973), "Transcendental Logarithmic Production Frontiers", *Review of Economics and Statistics* 55, 28-45.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1975), "Transcendental Logarithmic Utility Functions", American Economic Review 65, 367-383.

Cobb, C. and P.H. Douglas (1928), "A Theory of Production", *American Economic Review*, Supplement, 18, 139-165.

Denny, M. (1974), "The Relationship between Functional Forms for the Production System", *Canadian Journal of Economics* 7, 21-31.

Denny, M., M. Fuss and L. Waverman (1981), "The Measurement and Interpretation of Total Factor Productivity in Regulated Industries with an Application to Canadian Telecommunications", pp. 179–218 in *Productivity Measurement in Regulated Industries*, T.G. Cowing and R.E. Stevenson (eds.), New York: Academic Press.

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

Diewert, W.E. (1973), "Functional Forms for Profit and Transformation Functions", *Journal of Economic Theory* 6, 284-316.

Diewert, W.E. (1974a), "Applications of Duality Theory", pp. 106-171 in *Frontiers of Quantitative Economics*, Volume 2, M.D. Intriligator and D.A. Kendrick (eds.), Amsterdam: North-Holland.

Diewert, W.E. (1974b), "Functional Forms for Revenue and Factor Requirements Functions", *International Economic Review* 15, 119-130.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1978), "Hicks' Aggregation Theorem and the Existence of a Real Value Added Function", pp. 17-51 in *Production Economics: A Dual Approach to*

Theory and Applications, Volume 2, M. Fuss and D. McFadden, editors, North-Holland, Amsterdam.

Diewert, W.E. (1981), "The Comparative Statics of Industry Long Run Equilibrium", *Canadian Journal of Economics* 14, 78-92.

Diewert, W.E. (1993), "Duality Approaches to Microeconomic Theory", pp. 105-175 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.

Diewert, W.E. (2009), "The Aggregation of Capital over Vintages in a Model of Embodied Technical Progress", *Journal of Productivity Analysis* 32, 1-19.

Diewert, W.E. (2011), "Measuring Productivity in the Public Sector: Some Conceptual Problems", *Journal of Productivity Analysis* 36, 177-191.

Diewert, W.E. (2012), "The Measurement of Productivity in the Nonmarket Sector", *Journal of Productivity Analysis* 37, 217-229.

Diewert, W.E. (2018), "Productivity Measurement in the Public Sector", forthcoming in *The Oxford Handbook of Productivity Analysis*, E. Grifell-Tatjé, C.A.K. Lovell and R.C. Sickles (eds.), Oxford University Press.

Diewert, W.E. and R. Feenstra (2017), "Estimating the Benefits and Costs of New and Disappearing Products", Discussion Paper 17-10, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and K.J. Fox (2008), "On the Estimation of Returns to Scale, Technical Progress and Monopolistic Markups", *Journal of Econometrics* 145, 174–193.

Diewert, W.E. and K.J. Fox (2010), "Malmquist and Törnqvist Productivity Indexes: Returns to Scale and Technical Progress with Imperfect Competition", *Journal of Economics*, 101, 73-95.

Diewert, W.E. and K.J. Fox (2016), "Sunk Costs and the Measurement of Commercial Property Depreciation", *Canadian Journal of Economics* 49, 1340-1366.

Diewert, W.E. and K.J. Fox (2017), "The Difference Approach to Productivity Measurement and Exact Indicators", Discussion Paper 17-05, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and K.J. Fox (2018), "Decomposing Value Added Growth into Explanatory Factors", forthcoming in *The Oxford Handbook of Productivity Analysis*, E. Grifell-Tatjé, C.A.K. Lovell and R.C. Sickles (eds.), Oxford University Press.

Diewert, W.E. and R.J. Hill (2010), "Alternative Approaches to Index Number Theory", pp. 263-278 in *Price and Productivity Measurement*, W.E. Diewert, Bert M. Balk, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Victoria Canada: Trafford Press.

Diewert, W.E. and D. Lawrence (2002), "The Deadweight Costs of Capital Taxation in Australia", pp. 103-167 in *Efficiency in the Public Sector*, Kevin J. Fox (ed.), Boston: Kluwer Academic Publishers.

Diewert, E., D. Lawrence and J. Fallon (2009), *The Theory of Network Regulation in the Presence of Sunk Costs*, Technical report prepared for the New Zealand Commerce Commission. Available at: https://econ.sites.olt.ubc.ca/files/2013/06/pdf_paper_erwin-diewert-theory-network-regulation.pdf

Diewert, W.E. and N. Huang (2011), "Capitalizing R&D Expenditures", *Macroeconomic Dynamics* 15, 537-564.

Diewert, W.E. and C.J. Morrison (1986), "Adjusting Output and Productivity Indexes for Changes in the Terms of Trade", *The Economic Journal* 96, 659-679.

Diewert, W.E. and T.J. Wales (1987), "Flexible Functional Forms and Global Curvature Conditions", *Econometrica* 55, 43-68.

Diewert, W.E. and T.J. Wales (1988), "A Normalized Quadratic Semiflexible Functional Form", *Journal of Econometrics* 37, 327-342.

Diewert, W.E. and T.J. Wales (1992), "Quadratic Spline Models for Producer's Supply and Demand Functions", *International Economic Review* 33, 705-722.

Diewert, W.E. and T.J. Wales (1993), "Linear and Quadratic Spline Models for Consumer Demand Functions", *Canadian Journal of Economics* 26, 77-106.

Diewert, W.E. and A.D. Woodland (1977), "Frank Knight's Theorem in Linear Programming Revisited", *Econometrica* 45, 375-398.

Dixit, A. and V. Norman (1980), *Theory of International Trade: A Dual, General Equilibrium Approach*, Cambridge, UK: Cambridge University Press.

Feenstra, R.C. (1994), "New Product Varieties and the Measurement of International Prices", *American Economic Review* 84:1, 157-177.

Feenstra, R.C. (2004), *Advanced International Trade: Theory and Evidence*, Princeton N.J.: Princeton University Press.

Feenstra, R.C., R. Inklaar and M.P. Timmer, (2015) "The Next Generation of the Penn World Table", *American Economic Review* 105, 3150-3182.

Fenchel, W. (1953), "Convex Cones, Sets and Functions", Lecture Notes at Princeton University, Department of Mathematics, Princeton, N.J.

Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

Fox, K.J. (1998), "Non-Parametric Estimation of Technical Progress", *Journal of Productivity Analysis*, 10, 235–250.

Fox, K.J. and R.Q. Grafton (2000), "Nonparametric Estimation of Returns to Scale: Method and Application", *Canadian Journal of Agricultural Economics* 48, 341–354.

Gábór-Toth, E. and P. Vermeulen (2017), "The Relative Importance of Taste Shocks and Price Movements in the Variation of Cost-of-Living: Evidence from Scanner Data", paper presented at the 15th Meeting of the Ottawa Group, Eltville am Rhein, Germany.

Gale, D, V.L. Klee and R.T. Rockafellar (1968), "Convex Functions on Convex Polytopes", *Proceedings of the American Mathematical Society* 19, 867-873.

Gorman, W.M. (1968), "Measuring the Quantities of Fixed Factors", pp. 141-172 in *Value, Capital and Growth: Papers in Honour of Sir John Hicks*, J.N. Wolfe (ed.), Chicago: Aldine.

Hardy, G.H., J.E. Littlewood and G. Polya, (1934), *Inequalities*, Cambridge, England: Cambridge University Press.

Hicks, J.R. (1946), *Value and Capital*, Second Edition, Oxford: Clarendon Press.

Hotelling, H. (1932), "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions", *Journal of Political Economy* 40, 577-616.

Hotelling, H. (1935), "Demand Functions with Limited Budgets", *Econometrica* 3, 66-78.

Hottman, C.J., S.J. Redding and D.E. Weinstein (2016), "Quantifying the Sources of Firm Heterogeneity", *Quarterly Journal of Economics*, 1291–1364.

Inklaar, R. and W.E. Diewert (2016), "Measuring Industry Productivity and Cross-Country Convergence", *Journal of Econometrics* 191, 426–433.

Kohli, U.R.J. (1978), "A Gross National Product Function and the Derived Demand for Imports and Supply of Exports", *Canadian Journal of Economics* 11, 167-182.

Kohli, U. (1990), "Growth Accounting in the Open Economy: Parametric and Nonparametric Estimates", *Journal of Economic and Social Measurement* 16, 125-136.

Kohli, U. (1991), *Technology, Duality and Foreign Trade: The GNP Function Approach to Modelling Imports and Exports*, Ann Arbor, MI: University of Michigan Press.

Konüs, A.A. (1924), "The Problem of the True Index of the Cost of Living", translated in *Econometrica* 7, (1939), 10-29.

Konüs, A.A. and S.S. Byushgens (1926), "K probleme pokupatelnoi cili deneg", *Voprosi Konyunkturi* 2, 151-172.

Lawrence, D. and E. Diewert (2006), "Regulating Electricity Networks: The ABC of Setting X in New Zealand", pp. 207-241 in *Performance Measurement and Regulation of Network Utilities*, T. Coelli and D. Lawrence (eds.), Cheltenham: Edward Elgar Publishing.

Leontief, W.W. (1941), *The Structure of the American Economy 1919-1929*, Cambridge, MA: Harvard University Press.

McFadden, D. (1966), "Cost, Revenue and Profit Functions: A Cursory Review", IBER Working Paper No. 86, University of California, Berkeley.

McFadden, D. (1978), "Cost, Revenue and Profit Functions", pp. 3-109 in *Production Economics: A Dual Approach*, Volume 1, M. Fuss and D. McFadden (eds.), Amsterdam: North-Holland.

McKenzie, L.W. (1956-7), "Demand Theory without a Utility Index", *Review of Economic Studies* 24, 184-189.

Neary, J.P. (2004), "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income", *American Economic Review* 94, 1411-1428).

Pollak, R.A. (1969), "Conditional Demand Functions and Consumption Theory", *Quarterly Journal of Economics* 83, 60-78.

Rockafellar, R.T. (1970), *Convex Analysis*, Princeton, N.J.: Princeton University Press.

Samuelson, P.A. (1947), *Foundations of Economic Analysis,* Cambridge, MA: Harvard University Press.

Samuelson, P.A. (1953), "Prices of Factors and Goods in General Equilibrium", *Review of Economic Studies* 21, 1-20.

Samuelson, P.A. (1967), "The Monopolistic Competition Revolution", pp. 105-138 in *Monopolistic Competition Theory: Studies in Impact*, R.E. Kuenne (ed.), New York: John Wiley.

Shephard, R.W. (1953), *Cost and Production Functions*, Princeton N.J.: Princeton University Press.

Shephard, R.W. (1970), *Theory of Cost and Production Functions*, Princeton N.J.: Princeton University Press.

Uzawa, H. (1962), "Production Functions with Constant Elasticities of Substitution", *Review of Economic Studies* 29, 291-299.

Uzawa, H. (1964), "Duality Principles in the Theory of Cost and Production", *International Economic Review* 5, 291-299.

Walters, A.A. (1961), "Production and Cost Functions: An Econometric Survey", *Econometrica* 31, 1-66.

Wiley, D.E., W.H. Schmidt and W.J. Bramble (1973), "Studies of a Class of Covariance Structure Models", *Journal of the American Statistical Association* 68, 317-323.

Wold, H. (1944), "A Synthesis of Pure Demand Analysis; Part 3", *Skandinanvisk Aktuarietidskrift* 27, 69-120.

Wold, H. (1953), *Demand Analysis*, New York: John Wiley.

Woodland, A.D. (1982), *International Trade and Resource Allocation*. Amsterdam: North Holland.