2

Functional Form Specification

Functional form specification is a central aspect of empirical demand analysis.[1] Our primary objective in this chapter is to elucidate the relationships among particular parametric forms by showing that particular functional forms can be viewed as members of families or classes of functional forms. Our secondary objective is to demonstrate how particular forms can be constructed from a set of basic building blocks.

We emphasize classes of functional forms rather than treating particular parametric forms in isolation for three reasons. First, a reader's guide to currently used parametric forms would be obsolete before publication. The menu of parametric forms used in empirical demand analysis has grown rapidly over the last two decades and there is every reason to expect this rapid growth to continue. Second, focusing on classes of functional forms enables the reader to place in perspective not only currently used parametric forms but also new ones as they are introduced. Critical consideration of alternative parametric forms is essential both in planning one's own research and in evaluating the work of others. Our strategy—offering an analytical taxonomy—emphasizes the relationships among particular forms and thus increases the reader's awareness of alternative specifications. Third, viewing particular parametric forms against the background of the classes to which they belong enables us to recognize more easily and economically the limitations of particular functional forms. For example, suppose we specify and estimate a parametric functional form and find that our estimates imply that food and clothing are substitutes; to interpret this result, we must know whether our parametric functional form can exhibit complementarity. If it cannot, then our finding that food and clothing are substitutes was dictated by our functional form and reflects nothing about our data or economic reality. Focusing on the behavioral implications of classes of functional forms draws attention to the behavioral restrictions they imply.

Awareness of alternative specifications will not enable us to avoid making restrictive assumptions—restrictive parametric assumptions are

[1] This chapter draws on Howe, Pollak, and Wales [1979], Pollak [1969, 1971a, 1971b, 1972], and Pollak and Wales [1980].

inevitable in empirical demand analysis. Such awareness, however, will enable us to recognize the behavioral restrictions inherent in particular functional forms and, hence, to select appropriate forms for particular problems. As the previous sentence suggests, we do not believe that there is a single, "one-size-fits-all" functional form that is ideal for all applications. Instead, we believe that the characteristics that make a particular functional form suitable for one application may well make it inappropriate for another. For example, household budget data typically present the investigator with wide variation in observed levels of total expenditure but limited price variation. Time series data, on the other hand, typically offer less variation in expenditure and more variation in relative prices. Thus, it is not surprising that the parametric forms best suited for analyzing household budget data differ from those best suited for analyzing per capita time series data.

Our ordering of topics in this chapter is necessarily somewhat arbitrary. In Section 1 we discuss classes of demand systems defined in terms of the role of expenditure. In addition to the homothetic case, we discuss demand systems linear and quadratic in expenditure and share systems log-linear and log-quadratic in expenditure. In Section 2 we discuss demand systems generated by preference orderings satisfying various separability assumptions. We begin with direct additivity, indirect additivity, and generalized additive separability, a class that includes direct and indirect additivity. We then discuss weak and strong separability—the separability concepts appropriate for groups of goods. In Section 3 we discuss flexible functional forms, a class that includes the generalized Leontief and all of the translog forms including the "Almost Ideal Demand System" (AIDS) proposed by Deaton and Muellbauer. Because we did not discuss flexible functional forms in Chapter 1, we begin Section 3 with a detailed examination of the translog family.

1. EXPENDITURE SPECIFICATIONS

In this section we consider classes of demand systems in which expenditure (i.e., total expenditure or income) enters in simple ways. We begin by discussing expenditure proportionality, the theoretically important but empirically improbable case in which each good's expenditure elasticity is unity. We then discuss the class of demand systems in which the demand for each good is a linear function of expenditure, an empirically useful class that includes the LES. Continuing in this vein, we consider the class of demand systems in which the demand for each good is a quadratic function of expenditure; the empirical results we report in Chapters 6 and 7 rely heavily on members of this quadratic class.

We then consider an alternative direction for generalizing expenditure proportionality. Expenditure proportionality is equivalent to the requirement that the demand system, written in share form, is independent of the level of total expenditure. Working with the demand system in share form, we generalize this by considering the class of demand systems in which shares are linear functions of expenditure raised to a power or linear functions of the log of expenditure. We also consider a further generalization of the logarithmic class in which shares are quadratic in the log of expenditure.

1.1. Expenditure Proportionality

Expenditure proportionality plays an important role in economic theory, but has little practical relevance to empirical demand analysis. "Engel's Law," one of the great empirical regularities to emerge from nineteenth-century studies of household budget data, asserts that the budget share of food is smaller for rich than for poor households. Expenditure proportionality, which implies that the budget share of every good is independent of the household's total expenditure, thus contradicts Engel's Law. Despite its simplicity, the assumption of expenditure proportionality is too unrealistic to serve even as a useful first approximation for empirical demand analysis. Hence, demand systems exhibiting expenditure proportionality—especially the Cobb–Douglas, the Leontief, and the CES—are useful in empirical demand analysis primarily as building blocks that enter into the construction of more general demand systems.[2] The situation is different in the analysis of production where homothetic production functions, especially those exhibiting constant returns to scale, play a key role. The Cobb–Douglas, the Leontief, the CES, and the homothetic translog (which we discuss in Section 3) play key roles in the empirical analysis of production.

A demand system is said to exhibit expenditure proportionality if the demand for each good is proportional to expenditure

$$h^i(P, \mu) = B^i(P)\mu, \tag{1}$$

or, equivalently, if all expenditure elasticities are unity. If a demand system generated by a well-behaved utility function exhibits expenditure proportionality, then the demand functions must be of the form

$$h^i(P, \mu) = \frac{g_i(P)}{g(P)}\mu \tag{2}$$

where the function $g(P)$ is homogeneous of degree 1 and g_i denotes its partial derivative with respect to the ith price. We prove this result in Appendix A as part of the Gorman polar form theorem, which characterizes demand systems linear in expenditure. The demand functions (1)

[2] Expenditure proportionality plays a crucial role in the theory of aggregation over goods (see, for example, Blackorby, Primont, and Russell [1978]). Most of the standard, tractable, closed-form examples of globally defined demand systems exhibit expenditure proportionality.

are generated by a direct utility function that is homothetic to the origin, that is, a function that is an increasing transformation of a function homogeneous of degree 1:

(3) $$U(X) = T[V(X)]$$

where $T'(\cdot) > 0$ and $V(\lambda X) = \lambda V(X)$. Here, as elsewhere in the book, we sacrifice generality for tractability. In particular, we assume that demand functions and utility functions are differentiable enough to support calculus-based arguments. In characterizing increasing transformations (e.g., to define homotheticity or to investigate the class of direct or indirect utility functions representing a preference ordering) we confine ourselves to increasing transformations with strictly positive first derivatives, even though this excludes such strictly increasing transformations as $T(z) = z^3$.

The Cobb–Douglas and the CES are leading examples of demand systems exhibiting expenditure proportionality. The direct utility function for the Cobb–Douglas is given by

(4) $$U(X) = \sum a_k \log x_k, \qquad a_i > 0, \qquad \sum a_k = 1.$$

The direct utility function for the CES is

(5) $$\begin{aligned} U(X) &= -\sum a_k x_k^c, \qquad a_i > 0, \qquad c < 0 \\ U(X) &= \sum a_k x_k^c, \qquad a_i > 0, \qquad 0 < c < 1. \end{aligned}$$

The Cobb–Douglas ordinary demand functions are given by

(6) $$h^i(P, \mu) = \frac{a_i \mu}{p_i}$$

and the CES ordinary demand functions by

(7) $$h^i(P, \mu) = \frac{(p_i/a_i)^{1/(c-1)}\mu}{\sum p_k (p_k/a_k)^{1/(c-1)}}.$$

The indifference maps corresponding to these utility functions are identical to those of the Cobb–Douglas and CES production functions. The parameter c is related to the elasticity of substitution, σ, by $\sigma = 1/(1-c)$. It is easy to verify that when $c = 0$ the CES demand system reduces to the Cobb–Douglas. The Cobb–Douglas utility function is the limiting form of the CES as c approaches 0; Arrow, Chenery, Minhas, and Solow [1961], who introduced the CES production function, provide a clear discussion.

For most of the classes of demand systems we consider, direct utility function characterizations of the preferences that generate them either are unavailable or are uninformative in the sense that they fail to provide a transparent characterization of preferences. Indirect utility function characterizations, on the other hand, are widely available and play a central role in the analysis. In the case of expenditure proportionality, one can use Roy's identity to show that the demand system (2) is generated by the indirect utility function

$$\psi(P, \mu) = \frac{\mu}{g(P)}. \tag{8}$$

The indirect utility function corresponding to the Cobb–Douglas direct utility function (4) is given by (8) where

$$g(P) = \prod (p_k)^{a_k}. \tag{9}$$

The indirect utility function corresponding to the CES direct utility functions (5) is given by (8) where

$$g(P) = \left[\sum a_k^{-1/(c-1)} p_k^{c/(c-1)} \right]^{(c-1)/c}. \tag{10}$$

The limiting case of the CES direct utility function as c approaches minus infinity is the Leontief or fixed-coefficient utility function

$$U(X) = \min_k \left\{ \frac{x_k}{a_k} \right\}, \qquad a_i > 0, \tag{11}$$

as Arrow, Chenery, Minhas, and Solow [1961] show. The ordinary demand functions are given by

$$h^i(P, \mu) = \frac{a_i \mu}{\sum a_k p_k}. \tag{12}$$

Because the Leontief utility function is not differentiable, Lagrangian multipliers cannot be used to find the ordinary demand functions. A straightforward derivation relies on the fact that the ratios $\{x_i/a_i\}$ must all equal a common value. The indirect utility function corresponding to the Leontief case is given by (8) where

$$g(P) = \sum a_k p_k. \tag{13}$$

1.2. Demand Systems Linear in Expenditure

Consider demand systems linear in expenditure:

$$h^i(P, \mu) = C^i(P) + B^i(P)\mu. \tag{14}$$

Gorman [1961] has shown that any demand system that is linear in expenditure and theoretically plausible (i.e., consistent with utility maximization) must be of the form

$$h^i(P, \mu) = f_i(P) - \frac{g_i(P)}{g(P)} f(P) + \frac{g_i(P)}{g(P)} \mu \tag{15}$$

where f(P) and g(P) are functions homogeneous of degree 1. Because the argument can be used in a number of situations, we prove Gorman's result in Appendix A. If we require linearity to hold for all price–expenditure situations in which P and μ are strictly positive, then we are back to expenditure proportionality: the budget constraint implies $\sum p_k(f_k - fg_k/g) = 0$ and nonnegativity of consumption near 0 expenditure implies $f_i - fg_i/g \geqslant 0$. Hence, $f_i - fg_i/g = 0$ for all i. For purposes of empirical demand analysis it is usually sufficient to require linearity in a region of the price–expenditure space.

Using Roy's identity, we can easily verify that the class of demand systems, (15), is generated by an indirect utility function of the form

$$\psi(P, \mu) = \frac{\mu - f(P)}{g(P)} \tag{16}$$

which is known as the "Gorman polar form." The difficult part of Gorman's theorem or any characterization theorem is going in the other direction—showing that any theoretically plausible demand system linear in expenditure can be written in the form (15) and finding the class of indirect utility functions that generate the entire class of demand functions.

The LES,

$$h^i(P, \mu) = b_i - \frac{a_i}{p_i} \sum p_k b_k + \frac{a_i}{p_i} \mu, \tag{17}$$

the basis of our discussion in Chapter 1, is an example of a demand system linear in expenditure. It is generated by a Gorman polar form indirect utility function where $f(P) = \sum p_k b_k$ and $g(P) = \prod p_k^{a_k}$, $\sum a_k = 1$, so $f_i = b_i$ and $g_i/g = a_i/p_i$.

Although the class of direct utility functions yielding demand systems linear in expenditure has not been fully characterized, an important subclass consists of all direct utility functions of the form

$$U(X) = T[V(X - b)] \tag{18}$$

where $T'(\cdot) > 0$ and $V(\lambda X) = \lambda V(X)$. A function U(X) satisfying (18) is said to be homothetic to the point $(b_1, \ldots, b_n)$. The indifference curves of U are scaled up (or scaled down) versions of a single base indifference curve and the income–consumption curves radiate from the translated origin $(b_1, \ldots, b_n)$. The Gorman polar form corresponding to a direct utility function homothetic to the point $(b_1, \ldots, b_n)$ is given by (16), where $g(\cdot)$ is the dual of the function $V(\cdot)$ and f(P) is the linear function

$$f(P) = \sum p_k b_k. \tag{19}$$

This linearity restriction on the form of the function f(P) demonstrates that the class of direct utility functions homothetic to a point in the commodity space does not exhaust the class of direct preferences generating demand systems locally linear in expenditure.

The "trick" used in (18) to generate a new preference ordering from an old one by "translating" the origin from $0 = (0, \ldots, 0)$ to the point $b = (b_1, \ldots, b_n)$ in the commodity space is quite general and can be applied to any direct utility function, $U(X)$. Furthermore, duality enables us to perform an analogous trick with the indirect utility function: by translating the origin in the normalized price space, we generate another new preference ordering. In Section 2 we use this trick in conjunction with indirect additivity.

Consider the members of the "homothetic to a point" subclass corresponding to the direct utility functions we discussed in conjunction with expenditure proportionality. The LES utility function

$$(20) \quad U(X) = \sum a_k \log(x_k - b_k), \qquad a_i > 0, \qquad (x_i - b_i) > 0, \qquad \sum a_k = 1,$$

is a translation of the Cobb–Douglas (4). Translating the two CES forms, (5), yields

$$(21) \quad \begin{array}{llll} U(X) = -\sum a_k(x_k - b_k)^c, & a_i > 0, & (x_i - b_i) > 0, & c < 0, \\ U(X) = \sum a_k(x_k - b_k)^c, & a_i > 0, & (x_i - b_i) > 0, & 0 < c < 1. \end{array}$$

The corresponding demand functions are of the form

$$(22) \quad h^i(P, \mu) = b_i - \gamma^i(P) \sum p_k b_k + \gamma^i(P)\mu$$

where

$$(23) \quad \gamma^i(P) = \frac{(p_i/a_i)^{1/(c-1)}}{\sum p_k (p_k/a_k)^{1/(c-1)}}.$$

The indirect utility functions are given by (16) where $f(P)$ is given by (19) and $g(P)$ by

$$(24) \quad g(P) = \left[\sum a_k^{-1/(c-1)} p_k^{c/(c-1)} \right]^{(c-1)/c}.$$

Translating the Leontief direct utility function (11) yields

$$(25) \quad U(X) = \min_k \left\{ \frac{x_k - b_k}{a_k} \right\}, \qquad a_i > 0.$$

The corresponding demand functions are given by

$$(26) \quad h^i(P, \mu) = b_i - \frac{a_i}{\sum a_k p_k} \sum p_k b_k + \frac{a_i}{\sum p_k a_k} \mu.$$

These demand functions are generated by the indirect utility function (16) where $f(P)$ is given by (19) and $g(P)$ by (13).

Any quadratic direct utility function yields demand functions locally linear in expenditure. The additive quadratic, which goes back to Gossen (see Samuelson [1947, p. 93]), is a special case ($c = 2$) of the additive utility function

$$(27) \qquad U(X) = -\sum a_k(b_k - x_k)^c, \qquad a_i > 0, \qquad (b_i - x_i) > 0, \qquad c > 1,$$

which yields linear demand functions of the form (22). The indifference map of (27) is homothetic to the point $(b_1, \ldots, b_n)$, which is a "bliss point," and the income–consumption curves are straight lines that converge at this point (see Figure 1). The bliss point must lie in the first quadrant or no positive x's satisfy the condition $(b_i - x_i) > 0$, which is necessary for well-behaved preferences. For the additive direct quadratic the indifference curves are concentric circles centered at the bliss point $(b_1, \ldots, b_n)$. The utility function (27) is defined only in a box-like region of the commodity space southwest of the bliss point; for values of μ greater than $\sum p_k b_k$, the own-substitution terms implied by (27) are positive, in violation of regularity conditions.

The ordinary demand functions corresponding to (27), like those corresponding to (21) are given by (22) where $\gamma(P)$ is of the form (23). We refer to (21) and (27) as the "generalized CES class," although (27) is not a generalization of a well-behaved CES utility function (e.g., if $c = 2$ and if all of the translation parameters are 0, then the implied "CES indifference map" consists of concentric circles centered at the origin; these indifference curves have the "wrong" curvature).

For demand systems linear in expenditure, the marginal budget share of each good—that is, the fraction of an extra dollar of expenditure devoted to each good—is independent of expenditure. In the special case of expendi-

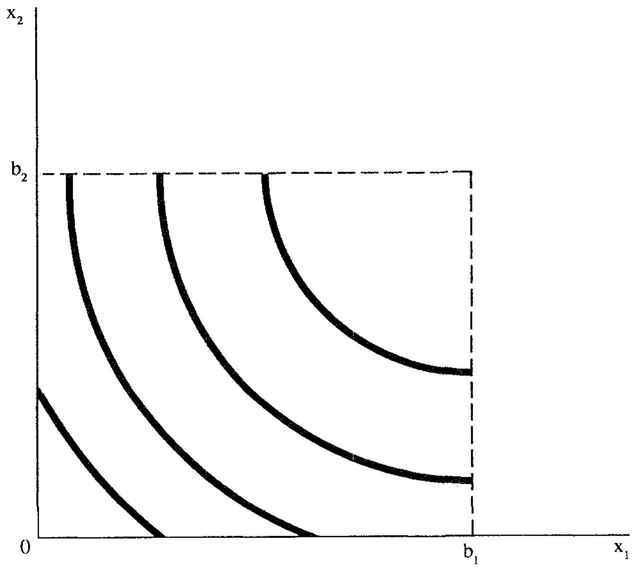

Figure 1 Additive direct quadratic indifference map

ture proportionality, marginal budget shares are not only independent of the level of expenditure but also equal to the average budget shares. For the LES the marginal budget shares are independent of prices as well as expenditure and are equal to the a's. In addition to the LES and the generalized CES, the class of demand systems linear in expenditure also includes the "linear translog" (LTL), a form we discuss in Section 3.

1.3. Demand Systems Quadratic in Expenditure

Casual empiricism suggests that the hypothesis that demand systems are linear in expenditure is too restrictive for the analysis of household budget data. Linearity implies that marginal budget shares are independent of the level of expenditure, so that rich households and poor households spend the same fraction of an extra dollar on each good. The implausibility of this hypothesis constitutes an argument for relaxing linearity so that it can be subjected to rigorous econometric testing. The class of demand systems quadratic in expenditure

$$h^i(P,\mu) = A^i(P)\mu^2 + B^i(P)\mu + C^i(P) \tag{28}$$

provides a natural generalization. van Daal and Merkies [1989], correcting a result of Howe, Pollak, and Wales [1979], show that if a quadratic demand system is theoretically plausible, then it is of the form

$$\begin{aligned} h^i(P,\mu) = {} & \frac{1}{\gamma}\left(\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right)\mu^2 + \left[\frac{g_i}{g} - \frac{2f}{\gamma}\left(\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right)\right]\mu \\ & + \frac{f^2}{\gamma}\left(\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right) - \frac{g_i}{g}f + f_i + \chi\left(\frac{g}{\gamma}\right)\frac{g^2}{\gamma}\left(\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right), \end{aligned} \tag{29a}$$

or, equivalently,

$$h^i(P,\mu) = \frac{1}{\gamma}\left(\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right)(\mu - f)^2 + \frac{g_i}{g}(\mu - f) + f_i + \chi\left(\frac{g}{\gamma}\right)\frac{g^2}{\gamma}\left(\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right), \tag{29b}$$

where the functions f, g, and γ are homogeneous of degree 1 and $\chi(\cdot)$ is a function of one variable. Howe, Pollak, and Wales [1979] failed in their attempt to characterize the entire class of demand systems quadratic in expenditure by implicitly assuming that $\chi(\cdot) = 0$. An alternative parameterization, instead of using the function $\gamma(P)$, uses the homogeneous of degree 1 function $\alpha(P)$ where $\alpha = g^2/\gamma$, so that the demand system (29b) becomes

$$h^i(P,\mu) = \frac{1}{g^2}\left(\alpha_i - \frac{g_i}{g}\alpha\right)(\mu - f)^2 + \frac{g_i}{g}(\mu - f) + f_i + \chi\left(\frac{\alpha}{g}\right)\left(\alpha_i - \frac{g_i}{g}\alpha\right). \tag{30}$$

Provided $\chi(\cdot) = 0$, these demand functions are generated by the indirect utility function

(31a) $$\psi(P, \mu) = -\frac{g(P)}{\mu - f(P)} - \frac{g(P)}{\gamma(P)},$$

or, equivalently,

(31b) $$\psi(P, \mu) = -\frac{g(P)}{\mu - f(P)} - \frac{\alpha(P)}{g(P)}.$$

The indirect utility function corresponding to the general case in which $\chi(\cdot) \neq 0$ is unknown. To facilitate comparing the indirect utility function (31) with that corresponding to demand systems linear in expenditure, we subject the Gorman polar form (16) to the increasing transformation $T^*(z) = -z^{-1}$ to obtain the equivalent indirect utility function

(32) $$\psi(P, \mu) = -\frac{g(P)}{\mu - f(P)}.$$

Adding a function homogeneous of degree 0 to (32) yields (31).

Specifying demand systems quadratic in expenditure is one way of relaxing the assumption that demand systems are linear in expenditure. Whether a particular quadratic generalization is worth the extra cost in degrees of freedom is in part an empirical question and in part a matter of taste and judgment. Since the quadratic hypothesis nests the linear one, classical hypothesis testing provides a framework for assessing whether a particular quadratic specification represents a statistically significant improvement over the linear one. But formal tests cannot resolve whether a statistically significant improvement is worth the extra cost in terms of difficulty of estimation and forgone opportunities to generalize in other directions.

We discuss two parametric specifications of demand systems quadratic in expenditure, the λ-QES and the Σ-QES. For both of these specifications $\chi(\cdot) = 0$ and the functions f(P) and g(P) are given by

(33) $$f(P) = \sum p_k b_k$$

and

(34) $$g(P) = \prod p_k^{a_k}, \qquad \sum a_k = 1.$$

For the λ-QES, the function α is given by

(35) $$\alpha(P) = \lambda g(P)^2 / \prod p_k^{c_k}, \qquad \sum c_k = 1;$$

for the Σ-QES, the function α is given by

(36) $$\alpha(P) = \sum p_k c_k.$$

These specifications, like any QES specification for which $\chi(\cdot) = 0$ and f and g are given by (33) and (34), reduce to the LES when $\alpha = 0$, or, equivalently, when $\gamma(P)$ is proportional to g(P).

1.4. Polynomial, Logarithmic, and Fractional Demand Systems

We open this section by discussing Muellbauer's characterization of two-term polynomial demand systems. We then introduce a general procedure for introducing constant terms (i.e., terms independent of expenditure) into theoretically plausible demand systems. We apply this procedure to Muellbauer's two-term polynomial demand system, obtaining polynomials with additional terms. We then discuss a remarkable theorem of Gorman that demonstrates the inherent limitations of polynomial demand systems. We conclude by discussing logarithmic and fractional demand systems.

The two-term polynomial demand system

$$h^i(P,\mu) = D^i(P)\mu^\sigma + B^i(P)\mu \tag{37}$$

is an attractive generalization of expenditure proportionality. Muellbauer [1975] designates this class, together with a logarithmic class that we discuss in the next section, by the acronym "PIGL" ("price independent generalized linearity").[3] He shows that theoretically plausible demand systems of the two-term polynomial form (37) are given by

$$h^i(P,\mu) = \left[\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right]\frac{\mu^\sigma}{\gamma^{\sigma-1}} + \frac{g_i}{g}\mu \tag{38}$$

where the functions g(P) and γ(P) are homogeneous of degree 1 and shows that these demand systems must be generated by the indirect utility function

$$\psi(P,\mu) = \frac{-\mu^{1-\sigma} - \gamma(P)^{1-\sigma}}{g(P)^{1-\sigma}}.\text{[4]} \tag{39}$$

We now propose a general procedure for introducing constant terms (i.e., terms independent of expenditure) into theoretically plausible demand systems. The trick is to replace the original demand system $\{\bar{h}^i(P,\mu)\}$ by the modified demand system $\{h^i(P,\mu)\}$ where

$$h^i(P,\mu) = f_i(P) + \bar{h}^i[P,\mu - f(P)] \tag{40}$$

and f(P) is a function homogeneous of degree 1. It is easy to verify that if the original system is generated by the indirect utility function $\bar{\psi}(P,\mu)$, then the modified system is generated by the indirect utility function

$$\psi(P,\mu) = \bar{\psi}[P,\mu - f(P)]. \tag{41}$$

---

[3] When $\sigma = 0$, two-term polynomial demand systems reduce to demand systems linear in expenditure.

[4] Muellbauer [1975] is primarily motivated by a concern with aggregation over individuals. Muellbauer works with demand systems in share form and uses the expenditure function rather than the indirect utility function. He also investigates the class of two-term demand systems exhibiting "generalized linearity," $h^i(P,\mu) = D^i(P)\phi(P,\mu) + B^i(P)\mu$.

Thus, if the original system is theoretically plausible, then the new system is also theoretically plausible, at least for values of f(P) near 0.

We illustrate this procedure with two examples. First, consider an original demand system exhibiting expenditure proportionality; applying our procedure, we obtain a demand system linear in expenditure. Replacing μ by $\mu - f(P)$ in the indirect utility function, we obtain the Gorman polar form. Second, consider the quadratic members of Muellbauer's PIGL class ($\sigma = 2$). It is easy to verify that modifying the quadratic PIGL indirect utility function

$$\bar{\psi}(P, \mu) = \frac{-\mu^{-1} - \gamma^{-1}}{g^{-1}} = -\frac{g}{\mu} - \frac{g}{\gamma} \tag{42}$$

by replacing μ by $\mu - f(P)$ yields the indirect utility function (31), a form that generates demand systems quadratic in expenditure.

It is evident that we can use this trick to generate higher degree polynomials from two-term polynomials. For example, suppose we begin with a two-term cubic demand system and replace μ by $\mu - f(P)$. The resulting system

$$h^i(P, \mu) = \left[\frac{g_i}{g} - \frac{\gamma_i}{\gamma}\right]\frac{(\mu - f)^3}{\gamma^2} + \frac{g_i}{g}(\mu - f) + f_i \tag{43}$$

when expanded, includes cubic, quadratic, and linear terms in μ as well as a term independent of μ. Notice, however, that the quadratic terms in the resulting demand system are proportional to the cubic terms (the factor of proportionality is $-3f$). The reader can verify that if we begin with a two-term quartic, then in the modified system both the cubic and the quadratic terms are proportional to the quartic terms. Thus, the higher order polynomial terms in demand systems generated in this way from two-term systems are severely restricted.

A remarkable theorem of Gorman's [1981] implies that such restrictions are not peculiar to the particular procedure we used to generalize two-term polynomial demand systems, but are an intrinsic feature of theoretically plausible polynomial demand systems. Gorman's theorem is complex, and we shall not state it in anything approaching its full generality. We shall, however, make plausible its implication that higher order terms in theoretically plausible polynomial demand systems must satisfy severe restrictions by considering the implications of the Slutsky symmetry conditions for third degree polynomial demand systems,

$$h^i(P, \mu) = D^i(P)\mu^3 + A^i(P)\mu^2 + B^i(P)\mu + C^i(P). \tag{44}$$

It is straightforward, although somewhat tedious, to calculate the Slutsky terms corresponding to this demand system

$$\frac{\partial h^i}{\partial p_j} + h^j \frac{\partial h^i}{\partial \mu} = \frac{\partial h^j}{\partial p_i} + h^i \frac{\partial h^j}{\partial \mu}. \tag{45}$$

These Slutksy terms are polynomials of degree 5 in μ and, because symmetry holds as an identity in the price–expenditure space, symmetry must hold for the coefficients of like powers of μ. (To see this, differentiate repeatedly with respect to μ and note that symmetry must hold for each of these derivatives). To illustrate our point, it suffices to calculate only that portion of the Slutsky term of degree 4 in μ. (The term of degree 5 in μ is given by $3D^jD^i$, so symmetry of this term imposes no restrictions on the coefficients of the demand system). The term of degree 4 is given by

$$2A^iD^j + 3A^jD^i \tag{46a}$$

or, equivalently,

$$[2A^iD^j + 2A^jD^i] + A^jD^i. \tag{46b}$$

Because the term in brackets is symmetric, symmetry of (46) implies

$$A^jD^i = A^iD^j \tag{47a}$$

or, equivalently,

$$A^j = \frac{A^i}{D^i}D^j. \tag{47b}$$

Hence, the coefficients of the quadratic terms (the A's in (44)) must be proportional to the coefficients of the cubic terms (the D's in (44)). The reader can verify that if we begin with a quartic demand system, then the Slutsky symmetry conditions imply that both the coefficients of the cubic terms and the coefficients of the quadratic terms must be proportional to the coefficients of the quartic terms. Gorman's theorem implies that increasing the number of terms in polynomial demand systems beyond three yields surprisingly little additional generality.

The PIGL class includes a logarithmic as well as a polynomial subclass. The logarithmic subclass, which Muellbauer calls PIGLOG, can be written as

$$h^i(P, \mu) = D^i(P)\mu \log \mu + B^i(P)\mu \tag{48}$$

or, in share form,

$$\omega^i(P, \mu) = \hat{D}^i(P) \log \mu + \hat{B}^i(P). \tag{49}$$

Muellbauer shows that all theoretically plausible PIGLOG demand systems can be written in the form

$$h^i(P, \mu) = \frac{g_i}{g}\mu - \frac{G_i}{G}[\log \mu - \log g]\mu \tag{50}$$

where g(P) is homogeneous of degree 1, G(P) is homogeneous of degree 0, and that the corresponding indirect utility function is given by

$$\psi(P, \mu) = G(P)[\log \mu - \log g(P)]. \tag{51}$$

The PIGLOG class includes not only expenditure proportionality but two functional forms that are widely used in empirical demand analysis, the log translog (log TL) and AIDS. We examine these two demand systems in detail in Section 3.

As Deaton [1981] points out, the PIGLOG class can be generalized in much the same way as Howe, Pollak, and Wales [1979] generalized the class of demand systems linear in expenditure to obtain demand systems quadratic in expenditure. Following Howe, Pollak, and Wales [1979] and van Daal and Merkies [1989], characterizing the theoretically plausible subclass should be straightforward. Because the empirical results we present in Chapter 6 indicate that the analysis of household budget data requires a demand system whose Engel curves have at least three independent parameters, we find the quadratic generalization of PIGLOG attractive.

In an important series of papers, Lewbel [1987a, 1987b, 1990] extends Gorman's theorem in several directions. Lewbel [1987a] introduces "fractional demand systems," which he defines as demand systems in which quantities or budget shares are proportional to an expression of the form: $A^i(P)\alpha(\mu) + B^i(P)\beta(\mu)$, where $\alpha(\cdot)$ and $\beta(\cdot)$ are differentiable functions of μ. The class of fractional demand systems contains demand systems linear in expenditure, some quadratic cases, PIGL, and PIGLOG. It also contains "translog-type" cases, including the "basic translog", which we discuss in Section 3. Lewbel obtains a closed-form characterization of the class of fractional demand systems and shows that, in addition to the cases just described, it contains three other cases which he calls "LOG2," "EXP," and "TAN."

Lewbel [1987b] investigates demand systems that are linear in expenditure and an arbitrary function of expenditure: $A^i(P)\alpha(\mu) + B^i(P)\mu + C^i(P)$. These "Gorman Engel Curves" are generalizations of the QES, PIGL, and PIGLOG; Lewbel establishes a closed-form characterization of the class.

Lewbel [1990] provides a closed-form characterization of "full rank demand systems." Gorman's theorem asserts that the coefficient matrix of a demand system that is linear in functions of expenditure is at most of rank three. Full rank demand systems are those with coefficient matrices of rank three. Lewbel proves that this class consists of PIGL, PIGLOG, a generalization of the quadratic, a quadratic logarithmic case, and a trigonometric case. Taken together, the Gorman and Lewbel results bind together the apparently disparate functional forms we have discussed in this section.

2. SEPARABILITY

In this section we discuss demand systems generated by separable preferences. Direct additivity and other separability assumptions, such as indirect additivity, restrict preferences by precluding certain types of specific interactions among goods. Additive preferences are most plausible when the goods correspond to broad aggregated categories (e.g., food, clothing, and recreation) rather than to narrow ones (e.g., bread, butter, and jam).

2.1. Direct Additivity

We say that a preference ordering is directly additive if it can be represented by a direct utility function of the form

$$U(X) = T[\sum u^k(x_k)] \tag{52}$$

where $T(\cdot) > 0$. The direct utility functions corresponding to the LES, (20), and the generalized CES, (21) and (27), exhibit direct additivity.

Direct additivity, as we have defined it, is ordinal; that is, it is a property of the preference ordering, not merely a property of the particular utility function chosen to represent preferences. When preferences exhibit direct additivity, the additive "canonical form" of the direct utility function

$$U(X) = \sum u^k(x_k) \tag{53}$$

has no better claim to being the "true" utility function than any other representation. Indeed, the phrase "true utility function" is foreign to modern demand analysis, which views utility functions as real valued representations of preference orderings—that is, as real valued functions defined on the commodity space that assign higher values to preferred commodity vectors.[5] For example, the preference ordering corresponding to the LES is equally well represented by

$$U(X) = \sum a_k \log(x_k - b_k), \qquad a_i > 0, \qquad (x_i - b_i) > 0, \qquad \sum a_k = 1, \tag{54a}$$

[5]In contrast to ordinal utility, "cardinal" utility does something more; for example, it may measure the intensity of an individual's preference for one basket of goods over another or compare the welfare that two individuals derive from the baskets of goods they consume. Replacing a utility function that measures intensity of preference or one that permits interpersonal comparisons by an increasing transformation of itself may alter the measure or the comparison. Hence, the set of admissible transformations of a cardinal utility function may be limited to some narrower class such as increasing linear transformations. None of this, however, is relevant to direct additivity because we have defined it as a property of the preference ordering.

or by

$$(54b) \qquad V(X) = \prod (x_k - b_k)^{a_k}, \qquad a_i > 0, \qquad (x_i - b_i) > 0, \qquad \sum a_k = 1,$$

or by any other increasing transformation of (54a).

Provided there are at least three goods —an assumption we maintain throughout this book—a preference ordering is directly additive if and only if the marginal rates of substitution involving each pair of goods depend only on the quantities of the two goods in the pair.[6] Equivalently, a preference ordering is directly additive if and only if the marginal rates of substitution involving each pair of goods are independent of the quantities of all other goods; that is,

$$(55) \qquad \frac{\partial}{\partial x_t}\left[\frac{U_i(X)}{U_j(X)}\right] = 0, \qquad t \neq i, j.$$

This necessary and sufficient condition for direct additivity can be reformulated as a condition on the preferences themselves. We begin with an example that violates direct additivity. Suppose there are three goods: bread (x_1), butter (x_2), and marmalade (x_3). An individual who lives by bread with butter and bread with marmalade, but not by bread alone, may prefer (10, 1, 10) to (5, 5, 10) and prefer (5, 5, 1) to (10, 1, 1). This pair of choices is incompatible with direct additivity. With direct additivity, if an individual prefers (10, 1, 10) to (5, 5, 10), then the individual will prefer $(10, 1, x_3)$ to $(5, 5, x_3)$ regardless of the common value of x_3. More generally, direct additivity implies that if an individual prefers $(x_1^a, x_2^a, \hat{x}_3)$ to $(x_1^b, x_2^b, \hat{x}_3)$ for a particular common value of x_3, say $\hat{x}_3$, then the individual prefers basket a to basket b for all common values of x_3. Direct additivity also implies analogous conditions when the common value good is x_1 or x_2 rather than x_3. Furthermore, if this "independence" condition holds for all partitions of the goods into two subsets—the common value good and the other two goods—then preferences exhibit direct additivity.[7] With three or more goods, preferences exhibit direct additivity if and only if, for all partitions of the goods into two subsets, an individual's preferences over vectors of goods that contain common values of the goods in one subset are independent of the levels of the goods in the common value subset.

With direct additivity the demand for each good is a function of the product of its own price and a single "index function" that depends on all prices and expenditure

$$(56a) \qquad h^i(P, \mu) = H^i[R(P, \mu)p_i].$$

[6]When there are only two goods, the conditions for and the implications of direct additivity are somewhat different, as Samuelson [1947, pp. 174–179] shows.

[7]Independence for one partition does not imply direct additivity but is compatible with an asymmetric preference structure such as $U(x_1, x_2, x_3) = V[V^*(x_1, x_2), x_3]$.

We restate this result in terms of normalized prices ($y_i = p_i/\mu$): with direct additivity the demand for each good is a function of the product of its own normalized price and an index function that depends on all normalized prices

(56b) $$h^i(Y) = H^i[S(Y)y_i].^8$$

The essence of this restriction is the existence of a single index function that appears in the demand function for each good. The LES is clearly of this form, since it can be written as

(57a) $$h^i(P, \mu) = b_i + \frac{a_i}{p_i}\left(\mu - \sum p_k b_k\right),$$

or, equivalently,

(57b) $$h^i(Y) = b_i + \frac{a_i}{y_i}\left(1 - \sum y_k b_k\right).$$

In the case of (57a) the index function is given by

(58a) $$R(P, \mu) = 1/[\mu - \sum p_k b_k].$$

In the case of (57b) the index function is given by

(58b) $$S(Y) = 1/[1 - \sum y_k b_k].$$

It is straightforward to show that direct additivity implies (56) and that the index function is the Lagrangian multiplier corresponding to the additive canonical form. The proof follows directly from the first order conditions for utility maximization. Using the additive canonical form, we write

(59) $$u^{i\prime}(x_i) = \lambda y_i, \qquad \sum y_k x_k = 1.$$

From these first order conditions we obtain

(60) $$x_i = H^i(\lambda y_i)$$

where $H^i(\cdot)$ is the inverse of the function $u^{i\prime}(\cdot)$ and λ is implicitly defined by the budget constraint

(61) $$\sum y_k H^k(\lambda y_k) = 1.$$

Hence, the demand functions are of the form (56b) and the index function

[8] Although it is sloppy notation, we use the same symbol to denote the functions $h^i(P, \mu)$ and $h^i(Y)$. It would be cleaner but more cumbersome to define a new function $h^{*i}(Y)$ by $h^{*i}(Y) = h^i(Y, 1)$, but our sloppiness should cause no confusion.

is the Lagrangian multiplier corresponding to the canonical form of the additive direct utility function.

The LES provides an instructive example. In terms of normalized prices the LES first order conditions are

$$\frac{a_i}{x_i - b_i} = \lambda y_i$$
$$\sum x_k y_k = 1. \tag{62}$$

Solving (62) for x_i yields

$$x_i = b_i + \frac{a_i}{\lambda y_i} \tag{63}$$

where λ is implicitly defined by the budget constraint

$$1 = \sum y_k x_k = \sum y_k b_k + \frac{1}{\lambda}\sum a_k = \sum y_k b_k + \frac{1}{\lambda}. \tag{64a}$$

With the LES we can solve explicitly for λ:

$$\lambda = \frac{1}{1 - \sum y_k b_k}, \tag{64b}$$

which is the index function appearing in Eq. (58b).

Returning to the general additive case, we next examine the restrictions on the derivatives of the demand functions that follow from the fact that both the prices of other goods and expenditure enter the demand for a particular good only through the index function. Differentiating (56a) with respect to p_j and μ yields:

$$\frac{\partial h^i}{\partial p_j} = H^{i\prime} R_j p_i, \qquad i \neq j \tag{65}$$

and

$$\frac{\partial h^i}{\partial \mu} = H^{i\prime} R_\mu p_i. \tag{66}$$

Provided (66) is not zero, we can eliminate the common factor between (65) and (66) to obtain

$$\frac{\partial h^i}{\partial p_j} = \theta^j(P, \mu)\frac{\partial h^i}{\partial \mu}, \qquad i \neq j \tag{67}$$

where the factor of proportionality θ^j is given by

$$\theta^j(P, \mu) = \frac{\partial R/\partial p_j}{\partial R/\partial \mu}. \tag{68}$$

That is, the change in the consumption of food induced by a change in the price of clothing is proportional to the change in the consumption of food induced by a change in expenditure; the factor of proportionality depends on the good whose price has changed, but not on the good whose quantity response we are considering. We can express (67) in ratio form as

$$\frac{\partial h^i/\partial p_j}{\partial h^t/\partial p_j} = \frac{\partial h^i/\partial \mu}{\partial h^t/\partial \mu}, \qquad j \neq i, t. \tag{69}$$

The analysis can be extended to the partial derivatives of the compensated demand functions. The ratio of the ijth substitution term to the product of the partial derivatives with respect to expenditure of the ordinary demand functions for goods i and j is independent of i and j. More precisely

$$\frac{(\partial h^i/\partial p_j) + h^j(\partial h^i/\partial \mu)}{(\partial h^i/\partial \mu)(\partial h^j/\partial \mu)} = \Gamma(P, \mu), \qquad i \neq j. \tag{70}$$

To show that this ratio is independent of i and j, we substitute for $\partial h^i/\partial p_j$ from (67) and observe that the ratio is independent of i. The parallel argument using the symmetric substitution term yields a similar expression that is independent of j. Hence, the ratio must be independent of both i and j as (70) asserts.

The assumption of direct additivity or "independent utilities" has a long history in economic analysis. Samuelson [1947, p. 93] attributes the general additive form to Jevons, while crediting Gossen with the additive quadratic utility function. The class of additive direct utility functions that yields demand functions exhibiting expenditure proportionality is known as the "Bergson family," in honor of Abram Bergson (Burk [1936]). In fact, Bergson investigated a somewhat different notion of independence and did not characterize the class that now bears his name. The Bergson family consists of the CES, its Cobb–Douglas limiting case, and sometimes (depending on an author's taste in regularity conditions) its Leontief and linear limiting cases.

The class of additive direct utility functions that yields demand functions locally linear in expenditure consists of the generalized CES, its two limiting cases, the LES and the translated Leontief, and additive exponential utility functions of the form

$$U(X) = -\sum \alpha_k e^{-\beta_k x_k}, \qquad \alpha_i > 0, \quad \beta_i > 0. \tag{71a}$$

It is convenient to rewrite (71a) as

$$U(X) = -\sum a_k e^{(b_k - x_k)/a_k} \tag{71b}$$

where $a_i = 1/\beta_i$ and $b_i = (\log \alpha_i\beta_i)/\beta_i$. The corresponding demand functions are

$$h^i(P, \mu) = b_i - \frac{a_i \sum p_k b_k}{\sum p_k a_k} + \frac{a_i \mu}{\sum p_k a_k} - a_i \log p_i + \frac{a_i \sum p_k a_k \log p_k}{\sum p_k a_k} \tag{72}$$

and the income–consumption curves are parallel straight lines. The indirect utility function is given by the Gorman polar form, (16), where

$$g(P) = \sum p_k a_k \tag{73}$$

and

$$f(P) = (\sum p_k a_k)(\log \sum p_k a_k) - \sum p_k b_k + \sum p_k a_k \log p_k. \tag{74}$$

Pollak [1971b], which characterized the class of additive utility functions yielding demand functions linear in expenditure, shows that the generalized CES and the exponential cases are related through the class of utility functions

$$U(X) = \sum \alpha_k (\beta_k + \delta_k x_k)^c. \tag{75}$$

It is easy to show that all admissible utility functions of this form yield demand systems locally linear in expenditure and that any admissible utility function of this form can be written in one of the three generalized CES forms. It is also clear that the exponential class is a limiting case of (75); more precisely,

$$\lim_{c \to -\infty} -\sum \alpha_k \left(1 + \frac{-\beta_k}{c} x_k\right)^c = -\sum \alpha_k e^{-\beta_k x_k}. \tag{76}$$

2.2. Indirect Additivity

We say that a preference ordering is indirectly additive if it can be represented by an indirect utility function of the form

$$\phi(Y) = T[\sum \phi^k(y_k)] \tag{77}$$

where $T(\cdot) > 0$ and the y's are normalized prices. Indirect additivity, like direct additivity, is an ordinal property. Applying Roy's identity, Eq. (19b) of Chapter 1 yields the ordinary demand functions:

$$h^i(Y) = \frac{\phi^{i'}(y_i)}{\sum y_k \phi^{k'}(y_k)} = \frac{\phi^{i'}(y_i)}{S(Y)}, \tag{78a}$$

or, in terms of unnormalized prices,

$$h^i(P, \mu) = \frac{\phi^{i'}(p_i/\mu)}{R(P, \mu)}, \tag{78b}$$

where the function R is defined in the obvious way. Indirect additivity thus implies that the demand for each good is a function of its own normalized price divided by an index function that depends on all normalized prices. Hence, the ratio of the demand functions for two goods depends on the normalized prices of the two goods but not on the index function:

$$\frac{h^i(Y)}{h^j(Y)} = \frac{\phi^{i'}(y_i)}{\phi^{j'}(y_j)}. \tag{79}$$

Although (78) and (79) are the most transparent ways to express the implications of indirect additivity, it is straightforward to establish the consequences of indirect additivity for the partial derivatives of the demand functions.

The "indirect addilog" (Houthakker [1960]) provides a good illustration of indirect additivity. The indirect addilog utility function is given by

$$\phi(Y) = \sum \frac{\alpha_k}{\beta_k + 1} y_k^{\beta_k + 1} \tag{80}$$

and the ordinary demand functions by

$$h^i(Y) = \frac{\alpha_i y_i^{\beta_i}}{\sum \alpha_k y_k^{\beta_k + 1}}. \tag{81}$$

Applying to the indirect addilog utility function the analogue of the translating trick applied to the direct utility function in Eq. (18), we replace each y_i by $y_i - \gamma_i$. This yields a new indirect utility function whose origin is the point $\gamma = (\gamma_1, \ldots, \gamma_n)$:

$$\phi(Y) = \sum \frac{\alpha_k}{\beta_k + 1} (y_k - \gamma_k)^{\beta_k + 1} \tag{82}$$

The corresponding ordinary demand functions are

$$h^i(Y) = \frac{\alpha_i (y_i - \gamma_i)^{\beta_i}}{\sum \alpha_k (y_k - \gamma_k)^{\beta_k + 1}} \tag{83}$$

The CES is a special case of the indirect addilog: when all of the β's are equal to a common value, then the indirect utility function (80) becomes

$$\phi(Y) = \sum \frac{\alpha_k}{\beta + 1} y_k^{\beta + 1} \tag{84}$$

and the ordinary demand functions (81) become

$$h^i(Y) = \frac{\alpha_i y_i^{\beta}}{\sum \alpha_k y_k^{\beta + 1}}. \tag{85}$$

It is easy to verify that (85) is the demand system corresponding to the CES, (7).

The CES demand system is exceptional in that the direct utility function and indirect utility function that generate it have a similar mathematical form. (The underlying notion of "self-dual preferences" was introduced by Houthakker [1965] and formalized in Pollak [1972].) In general, the direct and indirect utility functions corresponding to a particular class of preference orderings have different forms. Only a narrowly circumscribed class of preference orderings exhibit both direct and indirect additivity. Hicks [1969] and Samuelson [1969] show that this class consists of the

CES, its two limiting cases, the Cobb–Douglas and the Leontief, and the preference ordering corresponding to the direct utility function

(86) $$U(X) = u^1(x_1) + \sum_{k=2}^{n} a_k \log x_k.$$

With the exception of these cases, indirect utility functions corresponding to additive direct utility functions are not additive. The reader can verify, for example, that the indirect utility functions corresponding to the LES and the generalized CES are not additive.

2.3. Generalized Additive Separability

The crucial common feature of direct additivity and indirect additivity is that the implied demand for each good can be written as a function of the normalized price of that good and a single index function that appears in every demand function. Following Pollak [1972] we say that a demand system with this feature exhibits "generalized additive separability" (GAS):

(87) $$h^i(Y) = F^i[y_i, S(Y)].$$

Conditions on the derivatives and ratios of derivatives are easily derived from (87). The class of demand systems exhibiting GAS includes not only direct additivity and indirect additivity, but also the Fourgeaud–Nataf [1959] class of demand functions

(88) $$h^i(P, \mu) = F^i\left[\frac{p_i}{R(P)}, \frac{\mu}{R(P)}\right]$$

where $R(P)$ is a function homogeneous of degree 1. The form Houthakker [1965] calls the "self-dual addilog"

(89) $$h^i(Y) = \alpha_i y_i^{\beta_i} S(Y)^{(1-\delta)+\delta\beta_i},$$

where $S(Y)$ is defined implicitly by

(90) $$\sum \alpha_k y_k^{\beta_k+1} S(Y)^{(1-\delta)+\delta\beta_k} = 1,$$

also exhibits GAS. Using a result of Gorman, Pollak [1972] characterizes the direct and indirect utility functions corresponding to GAS.

2.4. Weak Separability and Strong Separability

A direct utility function, $U(x_1, \ldots, x_n)$ is said to exhibit "weak separability" if there exists a partition of the n goods into m subsets, m functions $V^r(X_r)$, and a function V such that

(91) $$U(X) = V[V^1(X_1), \ldots, V^m(X_m)]$$

where $m \geqslant 2$ and X_r is the vector of goods in the rth subset. It is convenient to use double subscripts to denote goods: the first subscript indicates the subset to which a good belongs, and the second indicates the particular good within the subset. Thus, x_{ri} denotes the ith good in the rth subset. We denote the number of goods in the rth subset by n_r. Thus, $X_r = (x_{r1}, \ldots, x_{rn_r})$ and $n_1 + n_2 + \cdots + n_m = n$.

A utility function is weakly separable if and only if the goods can be partitioned into subsets in such a way that every marginal rate of substitution involving two goods from the same subset depends only on the goods in that subset. If the utility function is weakly separable, it is easy to verify that this condition is satisfied. Leontief [1947a, 1947b] shows that this condition is also sufficient.

Weak separability has strong intuitive appeal. If we partition the goods into categories such as "food," clothing," and "recreation," it is tempting to assume that the utility function is weakly separable with respect to the subsets of this partition. (Strictly speaking, it is the set of indexes by which the commodities are identified—the integers from 1 to n—that are partitioned.) Similarly, if $U(X_1, \ldots, X_T)$ is an intertemporal utility function, where X_t denotes the vector of goods consumed in period t, it is again tempting to assume that the utility function exhibits weak separability, reflecting a type of intertemporal independence. The crucial assumption in both of these cases is that the individual's preference between two vectors of goods that differ only in the components of one subset, say "food," is independent of the identical nonfood components of the two vectors.

To establish the implications of weak separability for the ordinary demand functions, we make use of "conditional demand functions" introduced in Pollak [1969, 1971a]. Consider an individual whose preferences can be represented by a utility function $U(x_1, \ldots, x_n)$ that is not necessarily separable. If the individual's consumption of one good has been determined before he enters the market, we say that the good has been "preallocated." We assume that the individual is not allowed to sell any of his allotment of a preallocated good, and that he cannot buy more of it. For definiteness, suppose that the nth good is preallocated, while the remaining $n - 1$ goods are available on the market at prices over which the individual has no control, and that his expenditure on the available goods, μ_θ, is also predetermined. The individual is supposed to choose quantities of the first $n - 1$ goods so as to maximize $U(x_1, \ldots, x_n)$ subject to the "budget constraint"

$$\sum_{k=1}^{n-1} p_k x_k = \mu_\theta \tag{92}$$

and the additional constraint

$$x_n = \bar{x}_n \tag{93}$$

where $\bar{x}_n$ denotes his allotment of the nth good. Hence, his demand for the goods available on the market depends on the prices of these goods, total expenditure on them, and his allotment of the nth good; that is,

(94) $$x_i = g^{i \cdot n}(p_1, \ldots, p_{n-1}, \mu_\theta, x_n), \qquad i \neq n.$$

We call the function $g^{i \cdot n}(\cdot)$ the "conditional demand function" for the ith good. The second superscript, n, indicates that the nth good is preallocated, a terminology suggested by the analogy with conditional probability.

Conditional demand functions can also be defined when more than one good is preallocated. In general, a conditional demand function expresses the demand for a good available on the market as a function of three kinds of variables: the prices of all goods available on the market, total expenditure on these goods, and the quantities of the preallocated goods.

Formally, we partition the set of all commodities into two subsets, θ and $\bar{\theta}$. We assume that the goods in θ are available on the market, while those in $\bar{\theta}$ are preallocated; thus, if $j \in \theta$, then x_j is available on the market, while if $j \in \bar{\theta}$, then x_j is preallocated. We denote total expenditure on the goods available on the market by μ_θ. The individual is supposed to maximize $U(x_1, \ldots, x_n)$ subject to the "budget constraint"

(95) $$\sum_{k \in \theta} p_k x_k = \mu_\theta$$

and the additional constraints

(96) $$x_j = \bar{x}_j, \qquad j \in \bar{\theta}.$$

The demand for a good available on the market depends on the prices of the goods available on the market, total expenditure on them, and the quantities of the preallocated goods. Thus

(97) $$x_i = g^{i \cdot \bar{\theta}}(P_\theta, \mu_\theta, X_{\bar{\theta}}), \qquad i \in \theta,$$

where P_θ denotes the vector of prices of the goods available on the market and $X_{\bar{\theta}}$ denotes the vector of preallocated goods. The function $g^{i \cdot \bar{\theta}}$ is the conditional demand function for the ith good; the second superscript, $\bar{\theta}$, indicates that the goods in $\bar{\theta}$ are preallocated.

Now consider the relation between conditional demand functions and ordinary demand functions. Suppose that only the nth good is preallocated and that the individual's allotment of the nth good is precisely equal to the amount he would have purchased when facing prices P with expenditure μ. That is, $x_n = h^n(P, \mu)$. Suppose further that the amount he has to spend on the first $n - 1$ goods is precisely equal to the amount he would spend on these goods when facing prices P with expenditure μ: $\mu - p_n h^n(P, \mu)$. In this situation, the individual will purchase the same quantities of each of the goods available on the market as he would purchase if he faced prices P with expenditure μ and the nth good were not preallocated; that is

(98) $$h^i(P, \mu) = g^{i \cdot n}[p_1, \ldots, p_{n-1}, M^{\theta}(P, \mu), h^n(P, \mu)], \qquad i \neq n,$$

where

(99) $$M^{\theta}(P, \mu) = \mu - p_n h^n(P, \mu) = \sum_{k=1}^{n-1} p_k h^k(P, \mu).$$

This identity follows directly from the definitions of ordinary demand functions and conditional demand functions as solutions to related constrained maximization problems.

A similar result holds when more than one good is preallocated. Let $h^{\bar{\theta}}(P, \mu)$ denote the vector of ordinary demand functions for the preallocated goods. For example, if the last two goods are preallocated, we have $h^{\bar{\theta}}(P, \mu) = [h^{n-1}(P, \mu), h^n(P, \mu)]$. If all goods were available on the market, total expenditure on the goods in $\bar{\theta}$ would be $\sum_{k \in \bar{\theta}} p_k h^k(P, \mu)$ and total expenditure on the goods in θ would be

(100) $$M^{\theta}(P, \mu) = \mu - \sum_{k \in \bar{\theta}} p_k h^k(P, \mu) = \sum_{k \in \theta} p_k h^k(P, \mu).$$

By the argument used to establish (98) it follows that

(101) $$h^i(P, \mu) = g^{i \cdot \bar{\theta}}[P_{\theta}, M^{\theta}(P, \mu), h^{\bar{\theta}}(P, \mu)], \qquad i \in \theta,$$

where M^{θ} is defined by (100).

Conditional demand functions, introduced in Pollak [1969], can be used to analyze consumer behavior under rationing, to provide simple proofs of the Hicksian aggregation theorem and Samuelson's LeChatelier principle, and to decompose the cross-price derivatives of the ordinary demand functions. In this book, however, we use them only to examine the implications of separability.

We now examine the conditional demand functions corresponding to a weakly separable utility function. We assume that the goods in one subset are available on the market, while all other goods are preallocated. For definiteness, suppose that the goods available on the market are in subset r, and that subset r is "food." Formally, $tj \in \theta$ if $t = r$ and $tj \in \bar{\theta}$ if $t \neq r$. In accordance with our previous notation, μ_{θ} denotes total expenditure on food and P_{θ} the vector of food prices. The conditional demand functions are determined by maximizing the utility function (91) subject to the budget constraint

(102) $$\sum_{k=1}^{n_r} p_{rk} x_{rk} = \sum_{k \in \theta} p_k x_k = \mu_{\theta}$$

and the additional constraints

(103) $$x_{sj} = \bar{x}_{sj}, \qquad sj \in \bar{\theta}.$$

If we absorb the constraints (103) into the utility function we obtain

$$\text{(104)} \quad V[V^1(\bar{X}_1), \ldots, V^{r-1}(\bar{X}_{r-1}), V^r(X_r), V^{r+1}(\bar{X}_{r+1}), \ldots, V^m(\bar{X}_m)].$$

Clearly, the utility maximizing values of $(x_{r1}, \ldots, x_{rn_r})$ are independent of the preallocated goods: regardless of the levels of the preallocated goods, the individual has only to maximize $V^r(X_r)$ subject to (102). Hence, the conditional demand functions for the goods in θ are of the form

$$\text{(105)} \qquad g^{ri \cdot \bar{\theta}}(P_\theta, \mu_\theta, X_{\bar{\theta}}) = g^{ri \cdot \bar{\theta}}(P_\theta, \mu_\theta).$$

It follows from the general relation between ordinary demand functions and conditional demand functions (101) that with weak separability

$$\text{(106)} \qquad h^{ri}(P, \mu) = g^{ri \cdot \bar{\theta}}[P_\theta, M^\theta(P, \mu)]$$

where M^θ is defined by (100).

This result expresses the key implication of weak separability. It implies that the demand for a good in a particular subset (e.g., Swiss cheese in the food subset) can be expressed as a function of the prices of the goods in that subset and total expenditure on those goods (food prices and total expenditure on food). Although it would be more convenient if the demand for Swiss cheese depended only on food prices and total expenditure (μ), this is not what separability implies. Instead, separability implies that expenditure and the prices of goods outside the food subset enter the demand functions for food only through their effect on total expenditure on food. Thus, if we work with expenditure on food rather than total expenditure on all goods, we can ignore the prices of goods outside the food category.

Our characterization of the ordinary demand functions implied by weak separability leads directly to several conclusions about the effects of finite price and expenditure changes. Consider the change in the consumption of Swiss cheese when there is a change in the price of some nonfood item, say, shoes; we do not require that the change in the price of shoes be small. The change in the price of shoes will cause the individual to change his total expenditure on food (μ_θ) and the change in total expenditure on food will cause a change in Swiss cheese consumption. Now consider a change in the price of another nonfood item, say, tennis balls. Suppose that the effect on μ_θ of this price change is the same as the effect on μ_θ of the change in the price of shoes. Then the change in the price of tennis balls and the change in the price of shoes will have the same effects on the consumption of Swiss cheese. Similarly, if a change in expenditure has the same effect on μ_θ as the change in the price of shoes, then it will also have the same effect on Swiss cheese consumption. A similar result holds for simultaneous changes in expenditure and in the prices of several nonfood items.

The literature on separability has been more concerned with the partial derivatives of the demand functions than with the demand functions themselves. The implications of weak separability for the partial derivatives of the demand functions follow immediately from the results we have already established. Differentiating (106) with respect to p_{sj} and μ we obtain

(107) $$\frac{\partial h^{ri}}{\partial p_{sj}} = \frac{\partial g^{ri \cdot \bar{\theta}}}{\partial \mu_\theta} \frac{\partial M^\theta}{\partial p_{sj}}, \qquad s \neq r,$$

(108) $$\frac{\partial h^{ri}}{\partial \mu} = \frac{\partial g^{ri \cdot \bar{\theta}}}{\partial \mu_\theta} \frac{\partial M^\theta}{\partial \mu}.$$

That is, the change in the consumption of Swiss cheese caused by a change in the price of shoes (expenditure) is proportional to the change in total expenditure on food caused by the change in the price of shoes (expenditure). Provided $\partial h^{ri}/\partial p_{t\ell} \neq 0$, we may express (107) in ratio form as

(109) $$\frac{\partial h^{ri}/\partial p_{sj}}{\partial h^{ri}/\partial p_{t\ell}} = \frac{\partial M^\theta/\partial p_{sj}}{\partial M^\theta/\partial p_{t\ell}}, \qquad s, t \neq r.$$

Provided $\partial h^{ri}/\partial \mu \neq 0$, we can eliminate $\partial g^{ri \cdot \bar{\theta}}/\partial \mu$ between (107) and (108) and obtain

(110) $$\frac{\partial h^{ri}}{\partial p_{sj}} = \nu_r^{sj} \frac{\partial h^{ri}}{\partial \mu}, \qquad s \neq r,$$

where ν_r^{sj} is defined by

(111) $$\nu_r^{sj} = \frac{\partial M^\theta/\partial p_{sj}}{\partial M^\theta/\partial \mu}.$$

That is, the change in the consumption of Swiss cheese induced by a change in the price of shoes is proportional to the change in the consumption of Swiss cheese induced by a change in expenditure. The factor of proportionality is not a constant, but a function of all prices and expenditure. It is the same for all food items (Swiss cheese, roast beef), but it does depend on the good whose price has changed. We can express (110) in ratio form as

(112) $$\frac{\partial h^{ri}/\partial p_{sj}}{\partial h^{r\ell}/\partial p_{sj}} = \frac{\partial h^{ri}/\partial \mu}{\partial h^{r\ell}/\partial \mu}, \qquad s \neq r.$$

We now turn from weak separability to "strong separability." A direct utility function is said to exhibit strong separability if there exists a partition of the commodities into m subsets, m functions $V^r(X_r)$, and a function T, $T'(\cdot) > 0$, such that

(113) $$U(X) = T\left[\sum_{r=1}^{m} V^r(X_r)\right]$$

If $m \geqslant 3$, a utility function is strongly separable if and only if the goods can be partitioned into subsets in such a way that every marginal rate of substitution involving goods from different subsets depends only on the goods in those two subsets. Because a utility function that is strongly separable with m subsets is also weakly separable with m subsets, our weak separability results apply directly to strong separability. Hence we discuss only the additional implications of strong separability.

A weakly separable utility function is also called a "utility tree" and the subsets "branches." Strong separability is sometimes called "block additivity" and the subsets "blocks." This terminology is useful in discussing the additional implications of strong separability. If a utility function is a tree with m branches, in general we cannot combine two branches into a single branch and treat the utility function as a tree with m − 1 branches. That is, with weak separability the marginal rate of substitution between Swiss cheese and tennis balls can depend on the quantity of shoes. With block additivity, however, it is always permissible to treat the goods in two (or more) blocks as a single blocks. For example, to combine the goods in blocks s and t into a single block, we first renumber the blocks so that s = 1 and t = 2, and then note that (113) can be rewritten as

$$(114) \qquad U(X) = T\left[V^1(X_1) + V^2(X_2) + \sum_{r=3}^{m} V^r(X_r) \right].$$

Hence, if the utility function is strongly separable, we may write the ordinary demand functions as

$$(115) \qquad h^{ri}(P, \mu) = g^{ri \cdot \bar{\theta}}[P_r, P_t, M^{\theta}(P, \mu)],$$

where θ denotes the set of all goods in blocks r and t, and $\bar{\theta}$ the goods in the remaining blocks. The implications of strong separability that go beyond those of weak separability can easily be established by combining blocks into superblocks and recognizing that the utility function is weakly separable in those superblocks. (When there are only two blocks this technique does not yield any implications of strong separability that go beyond those of weak separability. Recall that direct additivity has different implications for cases with two goods and cases with three or more goods.)

To illustrate this superblock technique, consider the case in which $\bar{\theta}$ denotes the goods in one block and θ the goods in the remaining m − 1 blocks. For definiteness, suppose that the goods in block s are in $\bar{\theta}$ and the remaining goods are in θ. Then

$$(116) \qquad h^{ri}(P, \mu) = g^{ri \cdot \bar{\theta}}[P_{\theta}, M^{\theta}(P, \mu)], \qquad r \neq s.$$

If we take block s to be clothing, then (116) implies that the demand for Swiss cheese can be expressed as a function of the prices of all nonclothing goods and total expenditure on all goods other than clothing. It is not clear that it is more useful to write the demand for Swiss cheese this way than as a function of all food prices and total expenditure on food; if the utility function is strongly separable, however, then we have the option of writing the demand functions in either form and the choice between them must depend on the problem at hand. If we are primarily interested in the effect of various price and expenditure changes on Swiss cheese consumption, then it is probably more convenient to write the demand for Swiss cheese as a function of all food prices and total expenditure on food. If, on the other hand, we are primarily interested in the effect of a change in the price of shoes on the consumption of a variety of nonclothing goods (e.g., Swiss cheese, tennis balls, etc.), then it is more convenient to write the demands for these goods as functions of all nonclothing prices and total expenditure on goods outside the clothing category. The implications of (116) for finite price and expenditure changes should be clear from our earlier discussion of weak separability. The derivation of the implications for the partial derivatives follows the same general route and yields

$$\frac{\partial h^{ri}}{\partial p_{sj}} = \nu^{sj} \frac{\partial h^{ri}}{\partial \mu}, \qquad s \neq r \tag{117}$$

where ν^{sj} is defined by

$$\nu^{sj} = \frac{\partial M^{\theta}/\partial p_{sj}}{\partial M^{\theta}/\partial \mu}. \tag{118}$$

In contrast to the tree case, (110), with block additivity the factor of proportionality is independent of r.

Although we examined the implications of direct additivity in Section 2.1 we return to it here because of its relation to weak and strong separability. To avoid the exceptional two-good case, we assume $n \geqslant 3$. If an individual's utility function is strongly separable with blocks corresponding to "food," "clothing," etc., and if we form the Hicksian composite commodities corresponding to these blocks, then the utility function defined in terms of these composite commodities is additive. In describing the implications of additivity, we shall refer to the commodities as "food," "clothing," etc.

If a utility function is additive, then it is also strongly separable; for regardless of how the goods are partitioned into subsets, they will satisfy (113). This means that the results for weak and strong separability can be used to deduce the implications of additivity. Let θ and $\bar{\theta}$ be any partition of the goods into two subsets. Then the demand for the goods in θ can be written as a function of the prices of the goods in θ and total expenditure on these goods. This result holds for all possible partitions of the commodities into subsets and, thus, when the utility function is additive, we may use any partition of the goods that is convenient. For example, we can write the demand for food as a function of the price of food, the price of recreation, and total expenditure on food and recreation:

$$h^r(P, \mu) = g^{r \cdot \theta}[p_r, p_t, M^\theta(P, \mu)] \tag{119}$$

where $\theta = \{r, t\}$ and all other goods are in $\bar{\theta}$. Or, if we prefer, we can write the demand for food as a function of the prices of all goods except clothing and total expenditure on all goods other than clothing:

$$h^r(P, \mu) = g^{r \cdot \bar{\theta}}[P_\theta, M^\theta(P, \mu)], \tag{120}$$

where $\theta = \{1, \ldots, s-1, s+1, \ldots, n\}$ and $\bar{\theta} = \{s\}$. The implications of these results for finite price and expenditure changes are obvious and we shall not discuss them. The implications for the partial derivatives can be derived by differentiating (120) with respect to p_s and μ; ratios of these expressions yield results analogous to those obtained under weak and strong separability and provide an alternative derivation of (69).

In our discussion of additivity in Section 2.1 we derived an expression involving the partial derivatives of the compensated demand functions, (70). Under weak separability, we can obtain an analogous result: if a utility function is a tree, then there exist $(n^2 - n)/2$ functions $\{\Gamma^{rs}(P, \mu)\}$, $\Gamma^{rs}(P, \mu) = \Gamma^{sr}(P, \mu)$, such that

$$\frac{(\partial h^{ri}/\partial p_{sj}) + h^{sj}(\partial h^{ri}/\partial \mu)}{(\partial h^{ri}/\partial \mu)(\partial h^{sj}/\partial \mu)} = \Gamma^{rs}(P, \mu) = \Gamma^{sr}(P, \mu), \qquad s \neq r. \tag{121}$$

Goldman and Uzawa [1964] show that this condition is necessary and sufficient for weak separability; we shall prove only necessity. To show that this ratio is independent of i and j, we substitute from (110) into the left-hand side of (121) and obtain an expression that depends on sj and r, but is independent of i. The parallel argument using the symmetric substitution term yields a similar expression that is independent of j. Hence, the ratio must be independent of both i and j, although it does depend on both r and s. We denote the common value of these ratios by $\Gamma^{rs}(P, \mu) = \Gamma^{sr}(P, \mu)$.

If the utility function is strongly separable, the argument used in the weakly separable case to establish (121) can be used to show

$$\frac{(\partial h^{ri}/\partial p_{sj}) + h^{sj}(\partial h^{ri}/\partial \mu)}{(\partial h^{ri}/\partial \mu)(\partial h^{sj}/\partial \mu)} = \Gamma(P, \mu), \qquad s \neq r. \tag{122}$$

That is, instead of $(n^2 - n)/2$ functions $\{\Gamma^{rs}(P, \mu)\}$, there is only a single function, $\Gamma(P, \mu)$.

The "S-branch utility function" of Brown and Heien [1972] provides an instructive example of a specification suitable for empirical work. They postulate a strongly separable utility function in which the block utility functions are of the generalized CES form and the aggregator utility function is CES. For example, the generalized CES forms (21) can be written as

$$V^r(X_r) = \left[\sum_{k=1}^{n_r} a_{rk}(x_{rk} - b_{rk})^{c_r} \right]^{1/c_r}, \tag{123}$$

and the CES aggregator utility function as

$$V[V^1(X_1), \ldots, V^m(X_m)] = \left[\sum_{r=1}^{m} a_r(V^r(X_r))^c \right]^{1/c} \tag{124}$$

Letting $\sigma_r = 1/(c_r - 1)$ and $\sigma = 1/(c - 1)$, the corresponding demand system is given by

$$x_{ri} = h^{ri}(P_r, \mu_r) = b_{ri} - \gamma^{ri}(P_r) \sum p_{rk} b_{rk} + \gamma^{ri}(P_r)\mu_r, \tag{125}$$

where

$$\gamma^{ri}(P_r) = \frac{(p_{ri}/a_{ri})^{\sigma_r}}{\sum_{k=1}^{n_r} p_{rk}(p_{rk}/a_{rk})^{\sigma_r}} \tag{126}$$

and

$$\mu_r = \frac{a_r^{\sigma} \left[\sum_{k=1}^{n_r} a_{rk}^{\sigma_r} p_{rk}^{1-\sigma_r} \right]^{(1-\sigma)/(1-\sigma_r)}}{\sum_{s=1}^{n} a_s^{\sigma} \left(\sum_{k}^{n_s} a_{sk}^{\sigma_s} p_{sk}^{1-\sigma_s} \right)^{(1-\sigma)/(1-\sigma_s)}} \left(\mu - \sum_{s=1}^{m} \sum_{k=1}^{n_s} p_{sk} b_{sk} \right) + \sum_{k=1}^{n_r} p_{rk} b_{rk}. \tag{127}$$

The demand system is, of course, linear in expenditure. The expression for μ_r is most easily derived by substituting the indirect utility functions for the generalized CES into (124).

It is tempting to consider decomposing estimation of the S-branch into two stages, at the first stage estimating the parameters of the m conditional generalized CES demand systems corresponding to each block, and at the second stage estimating the m parameters of the CES aggregator utility function. Unfortunately, the validity of this two-stage approach to estimation depends on postulating an implausible stochastic structure that prevents disturbances associated with the demand functions for goods in one block from affecting the demand for goods in other blocks.

We conclude this section with a brief note on the literature. Leontief [1947a, 1947b] investigated the underlying mathematical structure of separability. Debreu [1960] provided an important characterization of additivity. The notions of weak and strong separability were developed in Strotz [1957, 1959] and Gorman [1959]. Goldman and Uzawa [1964] developed a characterization of separability in terms of the partial derivatives of the demand functions. Gorman [1968] provides a definitive discussion of separability concepts; even the reader who is not interested in the technical details will enjoy the subsequent exchange between Vind

[1971a, 1971b] and Gorman [1971a, 1971b]. Blackorby, Primont, and Russell [1978] provide a thorough discussion and rigorous analysis of separability.

3. FLEXIBLE FUNCTIONAL FORMS

Over the last decade flexible functional forms have come to play an increasingly important role in empirical demand analysis. The translogs are the most widely used family of flexible functional forms, and we begin this section by providing an extensive discussion of translog specifications. This marks a departure from our usual practice of emphasizing broad classes of demand systems rather than particular specifications. We deviate in this case not only because of the importance of translog specifications, but also because we think that translogs, like many other topics, are "best understood by staring hard at some non-trivial examples" (Halmos [1958, p. 42]). In Section 3.2 we use the translog family as the starting point for a more general discussion of flexible functional forms. Finally, in Section 3.3 we discuss the advantages and limitations of flexible specifications in empirical demand analysis.

3.1. The Translog Family

The translog, the most widely used family of flexible functional forms, provides a good introduction to the entire class. The "homothetic translog" (HTL), the simplest member of the translog family, is the best starting point. The HTL indirect utility function is given by

(128) $$\psi(P,\mu) = \log\mu - \sum \alpha_k \log p_k - \tfrac{1}{2}\sum_j \sum_k \beta_{kj} \log p_k \log p_j,$$

$$\beta_{ij} = \beta_{ji} \text{ for all i, j}, \qquad \sum \beta_{ki} = 0 \text{ for all i}, \qquad \sum \alpha_k = 1.$$

Recall our notational convention that a summation symbol without an index of summation indicates a sum over the index k. Hereafter, we shall use double summation symbols without indexes of summation, $\sum\sum$, to indicate sums over the indexes j and k; such double sums appear often in translogs and other flexible functional forms. Such forms often require symmetry restrictions on the elements of a coefficient matrix; hereafter, when we write $\beta_{ij} = \beta_{ji}$, we do so with the understanding that it holds for all i, j. Similarly, a requirement such as $\sum \beta_{ki} = 0$ is understood to hold for all i.

When working with translog demand systems, it is convenient to rewrite Roy's identity in share form:

(129) $$\omega^i(P,\mu) = -\frac{\partial\psi(P,\mu)/\partial\log p_i}{\partial\psi(P,\mu)/\partial\log\mu}.$$

It is also convenient to make use of the following mathematical fact, which the reader can easily verify: if $\beta_{kj} = \beta_{jk}$, then

$$\frac{\partial}{\partial \log p_i}\left[\frac{1}{2}\sum\sum \beta_{kj} \log p_k \log p_j\right] = \sum_j \beta_{ij} \log p_j. \tag{130}$$

Using (129) and (130), it is straightforward to show that the HTL share equations are given by

$$\omega^i(P, \mu) = \alpha_i + \sum_j \beta_{ij} \log p_j. \tag{131}$$

As its name indicates, the HTL corresponds to homothetic preferences and, hence, the corresponding demand system exhibits expenditure proportionality. Although this property makes the HTL uninteresting for empirical demand analysis, we discuss it here for three reasons. First, because the HTL is the simplest translog, it provides a good introduction to the entire family. Second, the HTL is a useful building block for constructing interesting consumer demand systems: it is, as we shall see, the common element that unites the translog family. Third, the HTL plays an important role in the analysis of producer behavior. The HTL indirect utility function, (128), can be solved explicitly for the log expenditure function or, in the terminology of production theory, the log cost function

$$\log \mu = \log q + \sum \alpha_k \log p_k + \frac{1}{2}\sum\sum \beta_{kj} \log p_k \log p_j, \tag{132}$$

where q is output and μ is total cost. This translog cost function, suitably generalized to allow nonconstant returns to scale and technical progress, is the most popular specification in the empirical analysis of cost functions and factor demand systems.

The HTL, like all members of the translog family, is a generalization of the Cobb–Douglas and reduces to it when all of the β's are 0. The α's of the HTL have a straightforward behavioral interpretation: when all prices are unity, the α's are equal to the expenditure shares. As we show below, the β's also can be given a straightforward behavioral interpretation.

The linear translog (LTL) indirect utility function is given by

$$\psi(P, \mu) = \log(\mu - \sum p_k b_k) - \sum \alpha_k \log p_k - \frac{1}{2}\sum\sum \beta_{kj} \log p_k \log p_j. \tag{133}$$

$$\beta_{ij} = \beta_{ji}, \qquad \sum \beta_{ki} = 0, \qquad \sum \alpha_k = 1.$$

The LTL is obtained by translating the origin of the HTL, a process that might loosely be described as introducing "committed quantities" into the homothetic system. The LTL share equations,

$$\omega^i(P, \mu) = \frac{p_i b_i}{\mu} + \left[\alpha_i + \sum_j \beta_{ij} \log p_j\right]\left[1 - \sum \frac{p_k}{\mu} b_k\right], \tag{134}$$

correspond to a demand system linear in expenditure. Both the LES and the HTL are special cases of the LTL. The LTL, the first translog form in the literature, was proposed by Lau and Mitchell (1971] and estimated by Manser [1976].

The log translog (log TL) indirect utility function is given by

$$\text{(135)} \qquad \psi(P,\mu) = -\sum \alpha_k \log(p_k/\mu) - \frac{1}{2}\sum\sum \beta_{kj} \log(p_k/\mu)\log(p_j/\mu),$$

$$\beta_{ij} = \beta_{ji}, \qquad \sum\sum \beta_{kj} = 0, \qquad \sum \alpha_k = 1,$$

and the corresponding share equations by

$$\text{(136)} \qquad \omega^i(P,\mu) = \frac{\alpha_i + \sum_j \beta_{ij}\log p_j - \log\mu \sum_j \beta_{ij}}{1 + \sum\sum \beta_{kj}\log p_j}.$$

Because the log TL share equations are linear in the log of expenditure, the log TL belongs to the PIGLOG class. The log TL has recently played a major role in demand system estimation; see, for example, Jorgenson [1990]. In the literature both the log TL—the term is ours—and a form we call the "basic translog" (BTL) are often referred to as "the translog," a practice virtually guaranteed to cause confusion.

The AIDS demand system of Deaton and Muellbauer [1980a]—the unfortunate acronym resulted from their decision to give the system a persuasive name ("Almost Ideal Demand System") rather than a descriptive one—is, like the log TL, a member of both the translog family and the PIGLOG class. We saw in Section 1.4 that the PIGLOG demand systems are generated by indirect utility functions of the form

$$\text{(137)} \qquad \psi(P,\mu) = G(P)[\log\mu - \log g(P)]$$

where $G(\lambda P) = G(P)$ and $g(\lambda P) = \lambda g(P)$ and that the share equations are given by

$$\text{(138)} \qquad \omega^i(P,\mu) = \frac{p_i g_i}{g} - \frac{p_i G_i}{G}[\log\mu - \log g].$$

The AIDS system is the special case in which

$$\text{(139)} \qquad G(P) = \prod p_k^{-\gamma_k}, \qquad \sum \gamma_k = 0$$

and

$$\text{(140)} \qquad \log g(P) = \alpha_0 + \sum \alpha_k \log p_k + \frac{1}{2}\sum\sum \beta_{kj}\log p_k \log p_j,$$

$$\beta_{ij} = \beta_{ji}, \qquad \sum \beta_{ki} = 0, \qquad \sum \alpha_k = 1.$$

It is easy to verify that the AIDS share equations are given by

$$\text{(141)} \qquad \omega^i(P,\mu) = \alpha_i + \sum \beta_{ki}\log p_k + \gamma_i[\log\mu - \log g].$$

The AIDS demand system reduces to the HTL when all of the γ's are 0.

The "basic translog" (BTL) indirect utility function, introduced by Christensen, Jorgenson, and Lau [1975], is given by

$$\psi(P,\mu) = -\sum \alpha_k \log(p_k/\mu) - \frac{1}{2}\sum\sum \beta_{kj} \log(p_k/\mu)\log(p_j/\mu), \tag{142}$$

$$\beta_{ij} = \beta_{ji}, \qquad \sum \alpha_k = 1.$$

It is straightforward to show that the BTL share equations are given by

$$\omega^i(P,\mu) = \frac{\alpha_i + \sum_j \beta_{ij} \log(p_j/\mu)}{1 + \sum\sum \beta_{kj} \log(p_j/\mu)}, \tag{143a}$$

or, equivalently,

$$\omega^i(P,\mu) = \frac{\alpha_i + \sum_j \beta_{ij} \log p_j - \log\mu \sum_j \beta_{ij}}{1 + \sum\sum \beta_{kj} \log p_j - B\log\mu}, \tag{143b}$$

where

$$B = \sum\sum \beta_{kj}. \tag{144}$$

(The normalization rule $\sum \alpha_k + \sum\sum \beta_{kj} = 1$ is sometimes more convenient than $\sum \alpha_k = 1$; when $\sum\sum \beta_{kj} = 0$, the two normalization rules are, of course, equivalent.)

The log TL is a special case of the BTL obtained by imposing the constraint

$$B = \sum\sum \beta_{kj} = 0, \tag{145}$$

so that expenditure drops out of the denominator of the share equation. The HTL is obtained from the BTL by imposing n parameter restrictions

$$\sum_j \beta_{ij} = 0, \qquad i = 1, \ldots, n. \tag{146}$$

With these n restrictions the expenditure term and the price terms drop out of the denominator of the share equations, and the expenditure term drops out of the numerator.

The "generalized translog" (GTL) is obtained by translating the BTL, in the same way that the LTL was obtained by translating the HTL. The indirect utility function is given by

$$\psi(P,\mu) = -\sum \alpha_k \log\left[p_k \Big/ \left(\mu - \sum_t p_t b_t\right)\right] - \frac{1}{2}\sum\sum \beta_{kj} \log\left[p_k \Big/ \left(\mu - \sum_t p_t b_t\right)\right] \log\left[p_j \Big/ \left(\mu - \sum_t p_t b_t\right)\right], \tag{147}$$

$$\beta_{ij} = \beta_{ji}, \qquad \sum \alpha_k = 1,$$

and the corresponding share system by

(148)

$$\omega^i(P,\mu)=\frac{b_i p_i}{\mu}+\left[1-\left(\sum p_k b_k\right)\Big/\mu\right]\frac{\alpha_i+\sum_j \beta_{ij}\log\left[p_j\Big/\left(\mu-\sum p_k b_k\right)\right]}{\sum \alpha_k+\sum\sum \beta_{jk}\log\left[p_j\Big/\left(\mu-\sum p_k b_k\right)\right]}.$$

The GTL includes as special cases both the BTL (when all of the b's are 0) and the LTL when ($\sum \beta_{ki}=0$ for all i).

The six forms we have discussed thus far are "indirect translog" forms: they correspond to preference orderings represented by indirect utility functions and they contain a double sum over the logs of prices. In addition to the indirect translogs, the literature on consumer and producer behavior contains a number of distinct translog forms whose common feature is that they somewhere contain a double sum over the logs of prices or quantities. A representative of the other branch of the translog family competing for the consumer behavior niche is the "direct basic translog," a preference ordering whose direct utility function is given by

$$U(X)=\sum a_k \log x_k+\frac{1}{2}\sum\sum b_{kj}\log x_k \log x_j, \tag{149}$$

$$b_{ij}=b_{ji}, \qquad \sum a_k=1.$$

Except in degenerate special cases, the direct translog and the indirect translog correspond to different preference specifications—hardly a surprise since, except in degenerate special cases, direct additivity and indirect additivity also correspond to distinct preference specifications. Because Roy's identity enables us to derive closed-form expressions for the share equations from indirect utility functions, indirect translogs are much more widely used than direct translogs.

Two translog forms are widely used in the study of producer behavior: the translog cost function, which we have already mentioned, and the translog profit function. Diewert [1982] provides extensive references to the literature. Although the translog profit function and the translog cost function have similar mathematical structures and are both flexible functional forms, they correspond to distinct specifications of the underlying technology.

It has been suggested that a major drawback of the translog, compared, for example, to the LES, is that the translog parameters lack a behavioral interpretation. To meet this objection, we now offer an interpretation of the translog β's. Before doing so, however, it is useful to present some preliminary results that enable us to express the Slutsky symmetry conditions in terms of elasticities and to express elasticities in terms of the derivatives of the share equations. The results in the following paragraph are general ones and do not assume that the underlying demand system is a translog.

As in Chapter 1, we denote the elasticity of the demand for the ith good with respect to the price of the jth good by E^i_j and its elasticity with respect to expenditure by E^i_μ

(150) $$E^i_j = \frac{p_j}{h^i}\frac{\partial h^i}{\partial p_j}$$

(151) $$E^i_\mu = \frac{\mu}{h^i}\frac{\partial h^i}{\partial \mu}.$$

It is easy to verify that the Slutsky symmetry condition

(152a) $$\frac{\partial h^i}{\partial p_j} + h^j\frac{\partial h^i}{\partial \mu} = \frac{\partial h^j}{\partial p_i} + h^i\frac{\partial h^j}{\partial \mu}$$

can be rewritten in terms of elasticities as

(152b) $$w_i E^i_j + w_i w_j E^i_\mu = w_j E^j_i + w_i w_j E^j_\mu$$

or, equivalently,

(152c) $$w_i E^i_j - w_i w_j E^j_\mu = w_j E^j_i - w_i w_j E^i_\mu.$$

The derivatives of the share equations with respect to the log price and log expenditure are closely related to elasticities. It is straightforward to verify that, for any demand system,

(153) $$\frac{\partial \omega^i(P,\mu)}{\partial \log p_i} = w_i(1 + E^i_i),$$

(154) $$\frac{\partial \omega^i(P,\mu)}{\partial \log p_j} = w_i E^i_j, \qquad i \neq j,$$

(155) $$\frac{\partial \omega^i(P,\mu)}{\partial \log \mu} = w_i E^i_\mu - w_i.$$

With these results in hand, we turn to the analysis of the translog.

We begin with the HTL because it yields strong simple results. These results are interesting because of the HTL's importance in empirical production analysis and because they generalize to other members of the translog family. Calculating the derivatives of the HTL share equations (131) with respect to $\log p_i$ and $\log p_j$, (153) and (154) imply

(156) $$w_i(1 + E^i_i) = \beta_{ii}$$

(157) $$w_i E^i_j = \beta_{ij}, \qquad i \neq j.$$

Because of the symmetry of the β's, (157) implies

(158) $$w_i E^i_j = w_j E^j_i.$$

Differentiating the HTL share equations with respect to log μ, (155) implies

$$E^i_\mu = E^j_\mu = 1, \tag{159}$$

a result that could have been derived directly from homotheticity. These results enable us to offer the following interpretation of the β's in the HTL: the off-diagonal β's are the product of the shares and the cross-price elasticities; the diagonal β's are the product of the shares and 1 plus the own-price elasticities. These relationships between the β's and the elasticities hold at all price–expenditure situations, not just when all prices and expenditure are unity.

The log TL and BTL yield more complex results which we derive in Appendix B. For the log TL

$$\frac{\beta_{ii}}{D(P,\mu)} = w_i[(1 + E^i_i) - w_i(E^i_\mu - 1)] \tag{160}$$

$$\frac{\beta_{ij}}{D(P,\mu)} = w_i E^i_j + w_i w_j - w_i w_j E^j_\mu, \qquad i \neq j, \tag{161}$$

where $D(P, \mu)$ is the expression that appears in the denominator of the log TL and BTL share equations:

$$D(P,\mu) = 1 + \sum\sum \beta_{kj} \log p_j - B \log \mu. \tag{162}$$

For the log TL, $B = 0$ so the final term is 0. When all prices and expenditure are unity, $D(P,\mu) = 1$ and (160) and (161) provide a behavioral interpretation of the β's in the log TL.

The BTL is more complicated. In Appendix B we define the function $\varepsilon(P,\mu)$ by $\varepsilon(P,\mu) = 2B/D(P,\mu)$ and show that

$$\varepsilon(P,\mu) = \frac{\partial(w_i E^i_\mu - w_i)/\partial \log \mu}{w_i E^i_\mu - w_i} = \frac{\partial^2 \omega^i(P,\mu)/\partial(\log \mu)^2}{\partial \omega^i(P,\mu)/\partial \log \mu} = \frac{2B}{D(P,\mu)}. \tag{163}$$

That is, the last two expressions in (163) are independent of i. We then show that

$$\frac{\beta_{ii}}{D(P,\mu)} = w_i[(1 + E^i_i) - w_i(E^i_\mu - 1)] + \tfrac{1}{2} w_i^2 \varepsilon(P,\mu) \tag{164}$$

$$\frac{\beta_{ij}}{D(P,\mu)} = w_i E^i_j + w_i w_j - w_i w_j E^j_\mu + \tfrac{1}{2} w_i w_j \varepsilon(P,\mu), \qquad i \neq j. \tag{165}$$

When all prices and expenditure are unity, these expressions provide a behavioral interpretation of the β's in the BTL, since the ε on the right hand side can be evaluated directly from the demand system using either of the last two expressions in (163).

3.2. Flexible Functional Forms

A demand system is said to be a "flexible functional form" if it is capable of providing a second order approximation to the behavior of any theoretically plausible demand system at a point in the price–expenditure space. More precisely, a flexible functional form can mimic not only the quantities demanded, the income derivatives, and the own-price derivatives, but also the cross-price derivatives at a particular point; equivalently, a flexible functional form can replicate not only the shares, the income elasticities, and the own-price elasticities, but also the cross-price elasticities at a specified price–expenditure situation. To understand the meaning of this definition, it is useful to begin by counting the number of shares and elasticities that a flexible function form must reproduce.

Consider an otherwise arbitrary theoretically plausible demand system and select a "point of approximation" in the price–expenditure space. At the point of approximation the demand system has n shares, n income elasticities, n own-price elasticities, and $n(n-1)$ cross-price elasticities. In a theoretically plausible demand system, however, not all of these n^2+2n values are independent. In particular,

- the shares must sum to unity, so only $n-1$ of them are independent;
- given the shares, only $n-1$ of the income elasticities are independent, since they must satisfy

$$\sum w_k E^k_\mu = 1, \tag{166}$$

  a relation easily derived by differentiating the budget constraint with respect to μ;
- given the shares and the income elasticities, only $n(n-1)/2$ of the cross-price elasticities are independent: given the "below diagonal" elasticities (i.e., E^i_j for $j < i$), the shares, and the income elasticities, we can infer the above diagonal elasticities using the Slutsky symmetry condition [in elasticity form, (152c)]:
- given the shares, the income elasticities, and the cross-price elasticities, all of the own-price elasticities are uniquely determined; we can infer them all, making use of the fact that the elasticities must satisfy

$$w_i + \sum w_k E^k_i = 0, \tag{167}$$

  a relation easily derived from the budget constraint by differentiating with respect to p_i.

Adding up these numbers, we find that there are at most $n(n-1)/2+2n-2$ independent shares and elasticities in a theoretically plausible demand system. We demonstrate below that a theoretically plausible demand system can in fact have this number of independent shares and elasticities.

A functional form is flexible with respect to the class of homothetic theoretically plausible demand systems if it is capable of providing a second order approximation to the behavior of any homothetic theoretically plausible demand system at a point in the price–expenditure space. Thus, a functional form that is flexible in this restricted sense need not be able to mimic arbitrary income elasticities. It must be possible for all of the income elasticities in such a system to equal unity, but it need not be possible for them to assume any other value. A slight modification of the argument used to show that a theoretically plausible demand system contains at most $n(n-1)/2 + 2n - 2$ independent shares and elasticities shows that a *homothetic* theoretically plausible demand system contains at most $n(n-1)/2 + n - 1$ independent shares and elasticities.

The HTL provides a transparent example of a functional form that is flexible with respect to the class of homothetic theoretically plausible demand systems and thus provides a useful starting point for discussing flexible functional forms. We begin by noting that the HTL contains $n(n-1)/2 + n - 1$ independent parameters: $n - 1$ α's (since they must sum to 1), n diagonal β's, and $n(n-1)/2$ off-diagonal β's, all subject to the n constraints of Eq. (146).

To show that the HTL can mimic the share and elasticity values of any homothetic theoretically plausible demand system, we begin with any such demand system and select a point of approximation. Without loss of generality, we can redefine the units in which the goods are measured so that in the redefined units all prices are unity at the point of approximation. (When we consider the nonhomothetic log TL and BTL, we shall require all normalized prices to be unity at the point of approximation; this amounts to requiring expenditure as well as all prices to be unity.) We denote the shares at the point of approximation by $\{\bar{w}_i\}$, the income elasticities that are, of course, unity by $\{\bar{E}^i_\mu\}$, and the price elasticities by $\{\bar{E}^i_j\}$. We now choose the HTL parameters as follows: let the α's equal the shares at the point of approximation; let the β's take the values implied by Eqs. (156) and (157):

$$\beta_{ii} = \bar{w}_i(1 + \bar{E}^i_i) \tag{168}$$

$$\beta_{ij} = \bar{w}_i\bar{E}^i_j, \qquad i \neq j. \tag{169}$$

We now show that the β's chosen in this way satisfy the symmetry condition, $\beta_{ij} = \beta_{ji}$, and Eq. (146). Symmetry of the β's requires (158), but this condition is automatically satisfied for any system exhibiting expenditure proportionality; it follows immediately from the Slutsky conditions (152b), in conjunction with (159). It is also easy to verify that

β's chosen in this way satisfy Eq. (146), since

(170) $$\sum_j \beta_{ij} = \sum_j \beta_{ji} = \sum_j \bar{w}_j \bar{E}_i^j + \bar{w}_i = 0$$

where the final step depends on Eq. (167).

With the HTL parameters chosen in this way, it is straightforward to verify that the HTL demand system will replicate the shares and the price elasticities (as well as the unitary income elasticities) of the original homothetic theoretically plausible demand system at the point of approximation.

The log TL is a flexible functional form, capable of mimicking arbitrary income elasticities $\{\bar{E}_\mu^i\}$ as well as arbitrary shares and the price elasticities of a theoretically plausible demand system at the point of approximation. Thus, is is not surprising that the log TL has $n(n-1)/2+2n-2$ independent parameters. (To see this, begin with the HTL count and note that the log TL contains $n-1$ additional parameters: the β's need not satisfy the n constraints (146), but only the single adding-up condition (145).) To show that the log TL is flexible, we choose the parameters as follows: let the α's equal the shares at the point of approximation; let the β's assume the values specified in Eqs. (160) and (161), (where $D = 1$):

(171) $$\beta_{ii} = \bar{w}_i + \bar{w}_i \bar{E}_i^i + \bar{w}_i^2 - \bar{w}_i^2 \bar{E}_\mu^i$$

(172) $$\beta_{ij} = \bar{w}_i \bar{E}_j^i + \bar{w}_i \bar{w}_j - \bar{w}_i \bar{w}_j \bar{E}_\mu^j, \qquad i \neq j.$$

We must show that the β's specified in this way satisfy the symmetry requirements, $\beta_{ij} = \beta_{ji}$, and that they satisfy the adding-up condition (145). Symmetry of the β's follows from the Slutsky symmetry conditions, most conveniently expressed in (152c). To show that the adding-up condition holds, we sum (171) and (172) over all j and find

(173) $$\sum_j \beta_{ij} = 2\bar{w}_i + \bar{w}_i \sum_j \bar{E}_j^i - \bar{w}_i \sum_j \bar{w}_j \bar{E}_\mu^j = \bar{w}_i - \bar{w}_i \bar{E}_\mu^i$$

where the final step uses (166) and the fact that the demand functions are homogeneous of degree 0 in prices and expenditure which implies

(174) $$\sum_j E_j^i + E_\mu^i = 0.$$

Summing (173) over i yields

(175) $$\sum_i \sum_j \beta_{ij} = 1 - \sum_i \bar{w}_i \bar{E}_\mu^i = 0$$

where the final step again makes use of (166). It is straightforward to verify that the log TL with the parameters chosen in this way mimics the behavior of the original demand system at the point of approximation.

The log TL is a parsimonious flexible functional form in the sense that it contains the minimum number of parameters required to achieve flexibility. Lau [1986, p. 1546] defines and discusses parsimony as well as other properties of flexible functional forms. The parsimony of the log TL follows immediately from a parameter count: a flexible functional form must contain at least $n(n-1)/2 + 2n - 2$ independent parameters, so any flexible form containing precisely this number of independent parameters is parsimonious.[9]

The BTL has $n(n-1)/2 + 2n - 1$ independent parameters, one more than the minimal number required for flexibility. Because the BTL is a generalization of the log TL, it is clearly a flexible functional form: given any set of shares and elasticities consistent with a theoretically plausible demand system, we can impose the adding-up constraint (145) on the β's and select the parameters so as to obtain the log TL that reproduces the given shares and elasticities. This procedure, however, fails to exploit the BTL's advantage over the log TL: the presence of an additional parameter, B. The key to exploiting this advantage is the observation that the subset of BTL forms corresponding to any value of B is flexible: given an arbitrary value of B, we can choose the remaining BTL parameters to mimic the shares, the income elasticities, and the price elasticities of any theoretically plausible demand system at the point of approximation. Empirical applications of the BTL rest on this fact, since there is no reason to expect the estimated value of B to be 0.

To prove that the BTL is flexible for an arbitrary value of B, $\bar{B}$, we select the BTL parameters as follows: let the α's equal the shares at the point of approximation; let the β's be given by the values corresponding to (164) and (165) where ε is equal to $2\bar{B}$ and D is unity:

$$\beta_{ii} = \bar{w}_i + \bar{w}_i\bar{E}^i_i + \bar{w}_i^2 - \bar{w}_i^2\bar{E}^i_\mu + \bar{w}_i^2\bar{B} \tag{176}$$

$$\beta_{ij} = \bar{w}_i\bar{E}^i_j + \bar{w}_i\bar{w}_j - \bar{w}_i\bar{w}_j\bar{E}^j_\mu + \bar{w}_i\bar{w}_j\bar{B}. \tag{177}$$

We must show that the β's specified in this way satisfy the symmetry requirements, $\beta_{ij} = \beta_{ji}$, and that they satisfy the adding-up condition. Symmetry of the β's follows from the Slutsky symmetry conditions, most conveniently expressed in (152c). To show that the BTL adding-up condition holds, we sum (176) and (177) over all j and find

$$\sum_j \beta_{ij} = \bar{w}_i - \bar{w}_i\bar{E}^i_\mu + \bar{w}_i\bar{B} \tag{178}$$

by an argument parallel to that used to derive (173). Summing (178) over i yields (144). It is straightforward to verify that the BTL with the parameters chosen in this way mimics the behavior of the original demand system at the point of approximation.

[9] These parameter counts refer to flexible functional forms for consumer demand systems. Counting parameters in producer demand system (i.e., demands for factors of production) is different because output is measurable while utility is not.

3.3. Advantages and Limitations

Flexible functional forms provide second order approximations to arbitrary twice differentiable functions. Instead of focusing on the ability of flexible functional forms to approximate preference orderings (i.e., to provide second order approximations to direct or indirect utility functions representing preferences), we have focused on their ability to provide first order approximations to theoretically plausible demand systems derived from preferences. The two approaches are fully consistent: since the partial derivatives of the demand functions are uniquely determined by the second partial derivatives of the utility function, the ability to approximate an arbitrary utility function implies the ability to approximate an arbitrary theoretically plausible demand system.[10]

We do not believe that the problem of modeling price effects has been solved by the development of flexible functional forms. Although the introduction of these forms constitutes an important expansion of the menu of specifications available for empirical research, they have yet to receive the critical scrutiny they require. In this section we discuss issues relevant to the assessment of flexible functional forms.

Suppose we want to estimate all of the elasticities of a demand system at a specified point of approximation in the price–expenditure space; suppose further that we can collect price–quantity data on which to base our estimates at any price–expenditure situations we select. Under these assumptions a plausible strategy is to specify a flexible functional form and estimate it using data corresponding to price–expenditure situations near the point of approximation. This strategy has seldom been implemented, presumably because economists seldom can select the price–expenditure situations at which data are collected.[11]

Viewed as local approximations, flexible functional forms avoid the restrictions inherent in nonflexible specifications; in doing so, however, they must either introduce additional parameters or imply other restrictions. For example, a nonflexible form with an equal number of parameters might be capable of approximating the second partial derivatives of the demand functions with respect to own-price and expenditure, at the cost of imposing restrictions on the cross-price derivatives. Thus, it is misleading to speak of a flexible functional form with K parameters as being "less restrictive" than a nonflexible specification with K independent parameters.

---

[10]To express the partial derivatives of the demand functions in terms of the second partials of the utility function, differentiate the first order conditions for utility maximization and solve using Cramer's rule, as in Samuelson [1947, pp. 100–102].

[11]There has been work based on experimental data, although little of it involves demand system estimation. The most prominent example of experimental data are those generated by time-of-day pricing experiments in electricity, by health care utilization experiments, and by negative income tax experiments.

Now suppose that we cannot select the price–expenditure situations at which data are collected, but instead must use price–quantity data corresponding to "natural" (i.e., uncontrolled, nonexperimental) price situations; for example, suppose our price data are deflators from the national product accounts or price indexes from the CPI. With such data the "local approximation" argument is irrelevant. To invoke it would require us to base our estimates on data from a neighborhood of the point of approximation and, without data from controlled experiments, the number of observations near the point of approximation is insufficient to permit estimation. To justify estimating a demand system using observations corresponding to price situations removed from the point of approximation, we must assume that the demand system we are estimating holds in a region large enough to include every data point we are using to estimate it. To motivate this assumption, it is often convenient to assume that it holds in a larger region, for example, in a connected region that includes the convex hull of the sample points. We call a demand system that holds in a region of the price–expenditure space an "exact" functional form in that region. Flexible functional forms, so appealing when viewed as local approximations, lose much of their appeal when interpreted as exact specifications.

When the number of goods is small, an attractive feature of flexible functional forms is that, compared to frequently estimated systems such as the LES and the generalized CES, they involve estimation of a large number of independent parameters. For example, with three goods the BTL involves eight parameters and the LES five. This fact, however, does not provide an argument for flexible functional forms because it applies equally to flexible and nonflexible functional forms involving the same number of parameters. For example, with three goods both the BTL and the QES contain eight parameters and it is not known whether the QES with three goods is flexible. Thus, although flexible functional forms have accustomed economists to estimating specifications involving many parameters, the availability of nonflexible multiparameter specifications precludes using this fact to justify the selection of a flexible functional form.

When the number of goods is large, the number of parameters in flexible functional forms, which increases with the square of the number of goods, becomes a liability. For example, with 15 goods the LES involves 29 parameters; the log TL, a parsimonious flexible functional form, involves 133. Empirical demand analysis has thus far avoided dealing with an unmanageable number of parameters by confining itself to systems involving a small number of goods.[12] Parsimonious or near parsimonious flexible functional forms are most attractive if one believes that the demand

[12]To retain at least some of the local approximation advantages of flexible functional forms while avoiding this proliferation of parameters, Diewert and Wales [1988] define and estimate "semiflexible" functional forms.

for a particular good is influenced as much by the price of any randomly selected good as by its own price or by expenditure. That is, a parsimonious flexible functional form corresponds to a "flat" or "uninformative" prior; it is this prior that justifies an allocation of parameters in which each specific cross-price effect, each own-price effect, and each expenditure effect receives equal weight. Because we believe that in most situations a randomly chosen specific cross-price effect is likely to be less important than the own-price or expenditure effect, we are skeptical of the allocation of parameters corresponding to parsimonious flexible functional forms. Flexible and nonflexible functional forms impose different restrictions on demand behavior, but in the absence of a "flat" prior, flexible functional forms may well be as restrictive as nonflexible specifications containing the same number of parameters. We emphasize this point because when flexible functional forms were first introduced, excessive claims were made for them. Although we dispute the excessive portions of these claims, the space we have devoted to flexible functional forms reflects our assessment of their importance.

When the data offer extensive variation in prices and expenditure and degrees of freedom are not a problem, this objection can be met by the use of specifications that are flexible and that include additional parameters to allow more "flexibility" in capturing expenditure and own-price effects. The generalized translog (GTL) of Pollak and Wales [1980] is an example of a specification of this sort.[13] When the available data do not permit estimation of additional parameters, we recommend sacrificing flexibility in order to obtain specifications better adapted to capturing own-price and expenditure effects.

It is often useful to view the choice of a functional form as a two-stage process. In the first stage, the number of independent parameters in the functional form is chosen; in the second, a particular form with the specified number of parameters is selected. Decomposing the choice in this way draws attention to the problem of allocating a fixed number of parameters so as to capture best the price and expenditure responses we want to estimate from a particular data set. The nature of the data set itself will almost certainly affect the choice of a functional form: household budget data usually offer substantial variation in expenditure levels and limited variation in relative prices. Time series data, on the other hand, usually offer substantially more variation in relative prices and less variation in expenditure. Thus, household budget data offer greater scope for estimating expenditure effects and should be analyzed using functional forms capable of reflecting them. Time series data, on the other hand, are likely to offer greater opportunities than household budget data for

[13] In the production context, the CES-translog of Pollak, Sickles, and Wales [1984] is another example.

estimating the specific cross-price effects that flexible functional forms are designed to model.

APPENDIX A: THE GORMAN POLAR FORM THEOREM

Theorem: Any theoretically plausible demand system linear in expenditure

$$h^i(P, \mu) = C^i(P) + B^i(P)\mu \tag{A1}$$

must be of the form

$$h^i(P, \mu) = f_i(P) - \frac{g_i(P)}{g(P)} f(P) + \frac{g_i(P)}{g(P)} \mu \tag{A2}$$

where f(P) and g(P) are functions homogeneous of degree 1. These demand functions are generated by the indirect utility function

$$\psi(P, \mu) = \frac{\mu - f(P)}{g(P)}. \tag{A3}$$

An obvious corollary is that any demand system exhibiting expenditure proportionality

$$h^i(P, \mu) = B^i(P)\mu \tag{A4}$$

must be of the form

$$h^i(P, \mu) = \frac{g_i(P)}{g(P)} \mu \tag{A5}$$

and is generated by the indirect utility function

$$\psi(P, \mu) = \frac{\mu}{g(P)}. \tag{A6}$$

Lemma: If the demand system (A1) is theoretically plausible, then

$$\sum p_k B^k = 1 \tag{A7}$$

$$\sum p_k C^k = 0 \tag{A8}$$

$$B^i_j = B^j_i \tag{A9}$$

$$C^i_j + C^j B^i = C^j_i + C^i B^j. \tag{A10}$$

These equalities hold as identities in P.

Proof of Lemma: The budget constraint

$$\mu = \mu \sum p_k B^k + \sum p_k C^k$$

holds as an identity, so terms in like powers of μ must be equal. This implies

(A7) and (A8). The Slutsky symmetry conditions imply that

$$\frac{\partial h^i}{\partial p_j} + h^j \frac{\partial h^i}{\partial \mu}$$

is symmetric in i and j (i.e., is equal to the corresponding expressions with i and j interchanged) and these equalities hold identically in (P, μ). Calculating this expression from (A1) we find

$$\frac{\partial h^i}{\partial p_j} + h^j \frac{\partial h^i}{\partial \mu} = (B^i_j + B^j B^i)\mu + (C^i_j + C^j B^i).$$

Since this holds as an identity in μ, terms in like powers of μ are equal to the corresponding terms with i and j interchanged. Equating terms in like powers of μ and cancelling where possible yields (A9) and (A10). QED

Proof of Theorem: The final assertion is easily verified using Roy's identity. The proof of the first part is broken down into two steps: first, we show that there exists a function f(P), homogeneous of degree 1, such that (A1) can be written as

$$\text{(A11)} \qquad h^i(P, \mu) = B^i(\mu - f) + f_i.$$

Second, we show that there exists a function g(P), homogeneous of degree 1, such that $g_i/g = B^i$. We assume throughout that the demand functions are differentiable enough to support the calculus arguments we employ and, hence, we establish our results only for this subclass.

In each part of the proof we appeal to a mathematical theorem on the existence of local solutions to systems of partial differential equations. The theorem guarantees the existence of a function $z = T(P)$, which satisfies a system of partial differential equations

$$\frac{\partial z}{\partial p_i} = \phi^i(P, z)$$

provided the symmetry conditions

$$\phi^i_j(P, z) + \phi^i_z(P, z)\phi^j(P, z) = \phi^j_i(P, z) + \phi^j_z(P, z)\phi^i(P, z)$$

hold (see Hurwicz and Uzawa [1971, Appendix]).

First, we show the existence of an f that satisfies

$$\text{(A12)} \qquad C^i = -B^i f + f_i$$

or, equivalently,

$$\text{(A13)} \qquad f_i = C^i + B^i f.$$

We define the functions $\phi^i(P, z)$ by

$$\phi^i(P, z) = C^i + B^i z.$$

By direct calculation

$$\phi^i_j + \phi^i_z \phi^j = C^i_j + B^i C^j + B^i_j z + B^i B^j z.$$

The fourth term is clearly symmetric. The symmetry of the first two terms is implied by (A10) and the symmetry of the third term by (A9). This establishes the existence of the function f(P).

To show that f(P) is homogeneous of degree 1, we multiply (A13) by p_i and sum over all goods to obtain

$$\sum p_k f_k = \sum p_k C^k + f \sum p_k B^k.$$

Making use of (A7) and (A8) we find

$$\sum p_k f_k = f.$$

The converse of Euler's theorem implies that f is homogeneous of degree 1.

Second, we show the existence of a function g(P) such that $g_i/g = B^i$. To do this, we define the function $\phi^i(P, z)$ by $\phi^i(P, z) = B^i(P)z$. Calculating $\phi^i_j + \phi^i_z \phi^j$, we find

$$\phi^i_j + \phi^i_z \phi^j = B^i_j z + B^i B^j z.$$

The second term is clearly symmetric and the first is symmetric in the light of (A9). Hence, there exists a function g(P) such that $g_i = B^i g$.

To show that g(P) is homogeneous of degree 1, we substitute g_i/g for B^i in (A7) and multiply by g to obtain

$$\sum p_k g_k = g.$$

The converse of Euler's theorem implies that g is homogeneous of degree 1.

QED

APPENDIX B: INTERPRETATION OF LOG TL AND BTL PARAMETERS

For the log TL, by differentiating the share equations (136) with respect to $\log p_i$, $\log p_j$, and $\log \mu$, we obtain equations analogous to (156), (157), and (159):

(B1) $$w_i(1 + E^i_i) = \frac{\beta_{ii}}{D} - w_i \frac{\sum \beta_{ki}}{D}$$

(B2) $$w_i E^i_j = \frac{\beta_{ij}}{D} - w_i \frac{\sum \beta_{kj}}{D}, \quad i \neq j,$$

(B3) $$w_i E^i_\mu - w_i = -\frac{\sum_j \beta_{ij}}{D},$$

where D is the denominator of the log TL share equation

(B4) $$D(P) = 1 + \sum\sum \beta_{kj} \log p_j.$$

Using the fact that the β's are symmetric, we substitute (B3) into (B1); solving for β_{ii}/D yields

$$\text{(B5)}\qquad \frac{\beta_{ii}}{D} = w_i[(1 + E^i_i) - w_i(E^i_\mu - 1)].$$

Similarly, substituting (B3) into (B2) and solving for β_{ij} yields

$$\text{(B6)}\qquad \frac{\beta_{ij}}{D} = w_i E^i_j + w_i w_j - w_i w_j E^j_\mu, \qquad i \neq j.$$

These expressions hold for all values of P and μ, but they simplify considerably when all prices and expenditure are unity because $D(1,\ldots,1,1) = 1$.

Differentiating the BTL share equations with respect to $\log p_i$ and $\log p_j$ yields expressions that look identical to (B1) and (B2), the expressions obtained from the log TL. They differ, however, in the role played by expenditure. In the log TL these expressions are independent of expenditure; in the BTL, expenditure enters through the function w_i in the numerator and the function D in the denominator:

$$\text{(B7)}\qquad D(P,\mu) = 1 + \sum\sum \beta_{kj} \log p_j - B \log \mu.$$

When all prices and expenditure are unity, these expressions simplify because $D = 1$. In the BTL case we differentiate the share equations with respect to $\log \mu$, obtaining

$$\text{(B8)}\qquad w_i E^i_\mu - w_i = -\frac{\sum_j \beta_{ij}}{D} + w_i \frac{B}{D}$$

where B is given by (144).

To clarify the meaning of the BTL parameters, we begin by differentiating (B8) with respect to $\log \mu$. After some rearranging this yields

$$\text{(B9)}\qquad \frac{\partial(w_i E^i_\mu - w_i)/\partial \log \mu}{w_i E^i_\mu - w_i} = \frac{2B}{D}.$$

This expression is independent of the choice of i, and we denote its common value by ε:

$$\text{(B10)}\qquad \varepsilon(P,\mu) = \frac{\partial(w_i E^i_\mu - w_i)/\partial \log \mu}{w_i E^i_\mu - w_i} = \frac{\partial^2 \omega^i(P,\mu)/\partial(\log \mu)^2}{\partial \omega^i(P,\mu)/\partial \log \mu} = \frac{2B}{D}.$$

In the case of the log TL, $B = 0$ and, hence, $\varepsilon = 0$. In the case of the BTL, these expressions are the same for all goods, and their common value is $2B/D$; when all prices and expenditure are unity, this reduces to 2B.

To obtain a behavioral interpretation of the β parameters in the BTL, we begin by rewriting (B8) as

$$\text{(B11)} \qquad \frac{\sum \beta_{ij}}{D} = w_i - w_i E^i_\mu + w_i \frac{B}{D} = w_i - w_i E^i_\mu + \tfrac{1}{2} w_i \varepsilon.$$

Relying on the symmetry of the β's and substituting this expression into (B1) and (B2) yields

$$\text{(B12)} \qquad \frac{\beta_{ii}}{D} = w_i[(1 + E^i_i) - w_i(E^i_\mu - 1)] + \tfrac{1}{2} w_i^2 \varepsilon$$

$$\text{(B13)} \qquad \frac{\beta_{ij}}{D} = w_i E^i_j + w_i w_j - w_i w_j E^j_\mu + \tfrac{1}{2} w_i w_j \varepsilon, \qquad i \neq j.$$

These are identical to Eqs. (164) and (165) in the text. As we noted there, these expressions simplify considerably when all prices and expenditure are unity, because in that case $D = 1$ and $\varepsilon = 2B$.