SPECIFICATION AND ESTIMATION OF CONSUMER DEMAND SYSTEMS WITH MANY BINDING NON-NEGATIVITY CONSTRAINTS

by

Lung-Fei Lee and Mark M. Pitt

Discussion Paper No. 236, October, 1986

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minn 55455

Abstract

This article presents coherent stochastic specifications for direct and indirect utility functions which result in computionally tractable demand systems subject to many binding non-negativity constraints. A seven commodity linear expenditure demand system for food is estimated using household consumption data.

Corresponding Address:

Lung-Fei Lee, 1035 Management and Economics, Department of Economics, University of Minnesota, 271 19th Avenue South, Minneapolis, Minnesota 55455.

SPECIFICATION AND ESTIMATION OF CONSUMER DEMAND SYSTEMS WITH MANY BINDING NON-NEGATIVITY CONSTRAINTS

by

Lung-Fei Lee and Mark M. Pitt(*)

1. Introduction

Household or individual microeconomic data offer important advantages for the analysis of consumer demand. Heterogeneous preferences associated with the age, sex, or educational attainments of consumers can be treated explicitly with micro data but cannot easily be incorporated into aggregate demand analysis. However, household budget data, which typically contain information on the consumption of disaggregated commodities, often demonstrate a significant proportion of observations for which expenditures on some goods are zero. Standard approaches of specifying and estimating demand systems are inappropriate in this case.

Recent papers by Wales and Woodland [18] and Lee and Pitt [11] have proposed methods for estimating demand systems with binding non-negativity constraints. The approach of Wales and Woodland is based upon the Kuhn-Tucker conditions associated with a stochastic direct utility function. They estimate a three-good model of the household demand for meat derived from a stochastic quadratic utility function. Lee and Pitt, taking the dual approach, begin with indirect utility functions and show how virtual price relationships can take the place of the Kuhn-Tucker conditions. They use this dual approach to estimate a three-good translog energy cost function with firm-level data (Lee and Pitt [8]). Both Wales and Woodland and Lee and Pitt have estimated only three-good examples because both of their stochastic specifications involve multiple numerical integrals which enormously complicate estimation for models with more than three goods. The generalization of these approaches to models with diaggregated commodities which are computationally tractable is a remaining practical issue.

This article reports several approaches to the stochastic specification of direct and indirect utility functions which imply likelihood functions which do not involve multiple integrals and may thus solve the computational issue. One of the proposed approaches is used to specify and estimate--with the method of maximum likelihood--a seven-good food demand system involving many zero demands.

Section 2 of this paper discusses the relation between the specification of stochastic terms in the utility function and model coherency with binding non-negativity constraints. Sections 3 and 4 propose stochastic specifications that are both computationally tractable and coherent. The results of estimating a seven-good demand system which utilizes one of our specifications are presented in Section 5. Our results are summarized in Section 6.

2. Random Preferences and Utility Maximization: Model Coherency

The econometric specification of consumer demand systems consistent with utility maximization requires attention to the manner in which stochastic terms are introduced. The common practice of simply appending additive error terms to demand equations derived from deterministic direct or indirect utility functions may result in a stochastic specification which is not compatible with utility maximization. Stochastic demand systems consistent with utility maximization are more likely to result if the random terms are introduced into the underlying direct or indirect utility functions, as in the approaches of Pollak and Wales [13], McFadden [12], Burtless and Hausman [4] and Wales and Woodland [18] among others. In the case of demand systems with binding non-negativity constraints, stochastic specifications also need to satisfy certain coherency conditions set forth below.

Let $U(x;\varepsilon)$ be a utility function with m commodity arguments $x_1, \cdots, x_m$, and a vector of stochastic terms ε which is fully known by each consumer but is stochastic from the econometricians perspective. It represents unobserved differences in consumers which affect their demands. The utility maximization problem of the consumer is

$$\text{(2.1)} \quad \max U(x;\varepsilon) \quad \text{subject to } v'x = 1 \quad \text{and } x_i \geq 0, \quad i = 1, \cdots, m$$

where $v = p/M$ is a vector of goods prices p normalized by income M. To guarantee unique solutions to (2.1), it is assumed that the utility function U is strictly increasing and strictly quasi-concave in x. It is furthermore assumed that U is continuously differentiable with respect to x.

Household-level consumption data with disaggregated commodities is often characterized by the nonconsumption of many goods by many consumers over a given reference period.[1] Let $x^* = (x_1^*, \cdots, x_m^*)$ be a vector of observed demand quantities ordered so that the first ℓ goods are not consumed and all remaining goods (indexed $\ell+1$ through m) are consumed. The optimal consumption vector x^* is characterized by the Kuhn-Tucker conditions:

$$(2.2)\quad \frac{\partial U(x^*;\varepsilon)}{\partial x_i} - \lambda v_i \leq 0 \qquad i = 1, \cdots, \ell$$

$$\frac{\partial U(x^*;\varepsilon)}{\partial x_j} - \lambda v_j = 0 \qquad j = \ell+1, \cdots, m$$

$$v'x^* = 1$$

where λ is the Lagrange multiplier corresponding to the budget constraint. The Kuhn-Tucker conditions can equivalently be expressed in terms of virtual prices as in Lee and Pitt [11]. Virtual prices are those prices which exactly support a vector of demands. In the case of a vector of demands having zero demands for some goods (such as x^*), virtual prices are derived by solving the consumer's problem (2.1) without the non-negativity constraints. Virtual prices can be written as

$$(2.3)\quad \xi_i = \frac{1}{\lambda}\,\frac{\partial U(x^*;\varepsilon)}{\partial x_i} \qquad i = 1, \cdots, m$$

Using virtual prices, the Kuhn-Tucker conditions can simply be rewritten as

$$(2.4)\quad \xi_i \leq v_i \qquad i = 1, \cdots, \ell$$

$$\xi_j = v_j \qquad j = \ell+1, \cdots, m$$

where $v'x = 1$. The use of these virtual price conditions is particularly advantageous in the dual approach to specifying demand systems subject to binding non-negativity constraints, as demonstrated in Lee and Pitt [11].

A functional form for U and a distribution for ε needs to be specified prior to parametric estimation. The likelihood function for x^* can then be derived using the relations (2.2) or (2.4) if the model is **coherent**. The structural equations (2.2) or (2.4) imply that the statistical model is essentially a simultaneous nonlinear equations model with multivariate limited dependent variables. Amemiya [1] and, subsequently, Gourieroux, Laffont and Monfort [6] have demonstrated that for these models certain coherency conditions must be satisfied to guarantee that the implied distribution functions for the endogenous variables are proper distributions. The coherency conditions are, first, that for every possible value of the ε vector, a unique vector of endogenous variables is generated by the structural equations; and, second, that for every possible vector of endogenous variables, there exists an ε vector that will generate them from the structural equations.

For our problem, the monotonicity and strict quasi-concavity properties of the utility function $U(x;\varepsilon)$ guarantees the existence of a unique demand vector for every ε vector. Satisfaction of the second coherency condition crucially depends on the manner in which the stochastic elements are introduced into the consumer's maximization problem. If there are m goods, typically at least $m-1$ stochastic terms are introduced. Less than $m-1$ stochastic terms may suggest some degree of determinism in consumer demand. More than $m-1$ stochastic terms may complicate the implied distributions for x^*. More important than the number of stochastic terms is the manner in which they are introduced. A given stochastic utility specification, functional form and vector of normalized prices may not be able to generate demand quantities which cover the entire simplex $\{x \mid v'x = 1, x \geq 0\}$. Thus it is possible that an observed outcome x^* is not realizable under this specification--violating the second coherency condition--and the likelihood for this observation will not exist. Satisfaction of the second coherency condition depends crucially on the functional form chosen for the utility function.

Attention to functional form is important not just to satisfy the coherency conditions. The presence of binding non-negativity constraints will complicate estimation if multivariate probabilities appear in the likelihood function. The following sections discuss how judicious choice of functional form and stochastic specification can allow for both coherency and tractable likelihood functions.

3. Stochastic Marginal Utilities: A Generalization

In their treatment of demand system estimation with binding non-negativity constraints, Wales and Woodland [18] begin with a random direct utility function. In their specification, each of the marginal utilities is composed of a deterministic component and an additive disturbance. The stochastic utility function is of the form $U(x;\varepsilon) = V(x) + \varepsilon' x$ where $V(x)$ is deterministic utility and ε is a vector of random terms of the same dimension as x.[2] They assume ε to be multivariate normal and impose $\sum_{i=1}^{m} \varepsilon_i = 0$ as a normalization rule. Strictly speaking, the assumption of multivariate normality is not compatible with the global monotonicity of the utility function. Monotonicity requires that $V(x)$ be strictly increasing and the random variables ε_i be all non-negative. A random utility specification consistent with monotonicity is

$$(3.1) \quad U(x;\varepsilon) = V(x) + \sum_{i=1}^{m} e^{\varepsilon_i} W_i(x_i)$$

where $V(x), W_i(x_i), i = 1, \cdots, m$ are strictly increasing and strictly concave functions. Computational tractability is achieved by assuming that the ε_i's are mutually independent and normal.[3] Since the utility function (3.1) is invariant under linear transformation, we can set a variance of ε (such as ε_m) to unity as a normalization.[4] There may be other useful normalizations as well. Treating $W_i(x_i)$ as a general function, rather than a linear function as in Wales and Woodland [18], generalizes the approach by allowing for more functional forms, in particular, the class of additive utility functions whose usefulness is demonstrated in the empirical section of this paper.

This random utility specification satisfies the model coherency requirements and implies a computationally tractable likelihood function for the model. Consider the general demand pattern with quantity vector $x^* = (0, \cdots, 0, x^*_{\ell+1}, \cdots, x^*_m)$ where the first ℓ goods are zero and the remaining goods are positive. The virtual prices at x^* are

$$(3.2)\quad \xi_i = \frac{v_m\left[\dfrac{\partial V(x^*)}{\partial x_i} + e^{\varepsilon_i}\dfrac{\partial W_i(x^*)}{\partial x_i}\right]}{\dfrac{\partial V(x^*)}{\partial x_m} + e^{\varepsilon_m}\dfrac{\partial W_m(x^*)}{\partial x_m}} \qquad i = 1, \cdots, m-1$$

The conditions (2.4) which characterize the optimality of x^* are equivalent to

$$(3.3)\quad e^{\varepsilon_i} \le \frac{\left[\dfrac{v_i}{v_m}\left[\dfrac{\partial V(x^*)}{\partial x_m} + e^{\varepsilon_m}\dfrac{\partial W_m(x^*)}{\partial x_m}\right] - \dfrac{\partial V(x^*)}{\partial x_i}\right]}{\dfrac{\partial W_i(x^*)}{\partial x_i}} \qquad i = 1, \cdots, \ell$$

and

$$(3.4)\quad e^{\varepsilon_j} = \frac{\left[\dfrac{v_j}{v_m}\left[\dfrac{\partial V(x^*)}{\partial x_m} + e^{\varepsilon_m}\dfrac{\partial W_m(x^*)}{\partial x_m}\right] - \dfrac{\partial V(x^*)}{\partial x_j}\right]}{\dfrac{\partial W_j(x^*)}{\partial x_j}} \qquad j = \ell+1, \cdots, m-1.$$

Given x^*, the range of values of ε_m which satisfy the second coherency requirement will be determined by the inequalities

$$\frac{v_i}{v_m}\left[\frac{\partial V(x^*)}{\partial x_m} + e^{\varepsilon_m}\frac{\partial W_m(x^*)}{\partial x_m}\right] - \frac{\partial V(x^*)}{\partial x_i} > 0, \qquad i = 1, \cdots, m-1$$

The feasible values of ε_m are, therefore, determined by

$$(3.5)\quad \varepsilon_m > \ln\left[\max\left\{0, \frac{\left[\dfrac{v_m}{v_i}\dfrac{\partial V(x^*)}{\partial x_i} - \dfrac{\partial V(x^*)}{\partial x_m}\right]}{\dfrac{\partial W_m(x^*)}{\partial x_m}} \quad \text{for } i = 1, \cdots, m-1\right\}\right]$$

where $\ln 0 = -\infty$ is understood as a conventional rule. Denote

$$(3.6) \quad \varepsilon_i(x^*,\varepsilon_m) = \ln\left[\frac{v_i}{v_m}\left[\frac{\partial V(x^*)}{\partial x_m} + e^{\varepsilon_m}\frac{\partial W_m(x^*)}{\partial x_m}\right] - \frac{\partial V(x^*)}{\partial x_i}\right] - \ln\frac{\partial W_i(x^*)}{\partial x_i} \qquad i = 1,\cdots,m-1$$

for ε_m satisfying (3.5) and let $R(x^*)$ be the expression on the right hand side of (3.5). As $x_m^* = (1 - v_{\ell+1}x_{\ell+1}^* - \cdots - v_{m-1}x_{m-1}^*)/v_m$, given the distributions $\varepsilon_{\ell+1},\cdots,\varepsilon_{m-1}$ conditional on ε_m, the relations (3.4) imply a conditional distribution for $x_{\ell+1}^*,\cdots,x_{m-1}^*$. Let f_i and F_i be, respectively, the density function and the distribution function of ε_i. Conditional on ε_m, the joint density function for $x_{\ell+1}^*,\cdots,x_{m-1}^*$ is

$$\prod_{i=\ell+1}^{m-1} f_i\left[\varepsilon_i(x^*,\varepsilon_m)\right] \, |J(x^*,\varepsilon_m)|$$

where the Jacobian $J(x^*,\varepsilon_m)$ is the determinant of the matrix

$$\begin{bmatrix} \dfrac{\partial\varepsilon_{\ell+1}(x^*,\varepsilon_m)}{\partial x_{\ell+1}^*} & \cdots & \dfrac{\partial\varepsilon_{\ell+1}(x^*,\varepsilon_m)}{\partial x_{m-1}^*} \\ \vdots & & \vdots \\ \dfrac{\partial\varepsilon_m(x^*,\varepsilon_m)}{\partial x_{\ell+1}^*} & \cdots & \dfrac{\partial\varepsilon_{m-1}(x^*,\varepsilon_m)}{\partial x_{m-1}^*} \end{bmatrix}$$

Since $\varepsilon_1,\cdots,\varepsilon_m$ are mutually independent, the conditional probability, conditional on $\varepsilon_{\ell+1},\cdots,\varepsilon_m$ (or equivalently, $x_{\ell+1}^*,\cdots,x_{m-1}^*,\varepsilon_m$), that the inequality conditions (3.3) hold is

$$\prod_{i=1}^{\ell} F_i(\varepsilon_i(x^*,\varepsilon_m)).$$

Hence the likelihood function $L(x^*)$ for the sample observation is

$$(3.7) \quad L(x^*) = \int_{R(x^*)}^{\infty} \prod_{i=1}^{\ell} F_i(\varepsilon_i(x^*,\varepsilon_m)) \cdot \prod_{i=\ell+1}^{m-1} f_i(\varepsilon_i(x^*,\varepsilon_m)) \cdot |J(x^*,\varepsilon_m)| f_m(\varepsilon_m)\, d\,\varepsilon_m.$$

With an independent sample of size N, the likelihood function for the whole sample will be

$$(3.8)\quad L(x_1^*, \cdots, x_N^*) = \prod_{i=1}^{N} L(x_i^*).$$

The likelihood function is computationally tractable as it involves effectively a single integral for each sample observation. In the empirical section of this paper, we apply this specification to the estimation of a linear expenditure system.

4. Random Preferences: Translating and Scaling

The stochastic elements in our utility function may represent unobservable characteristics of consumers. Translating and scaling are two familiar methods for introducing consumer characteristics into demand systems (see Pollak and Wales [14]). Below we consider these procedures in the context of both an additive utility function and an indirect utility function with binding non-negativity constraints.

a. Translating

The translating procedure consists of appending additive terms to the consumption vector elements in the utility function. In the case of binding non-negativity constraints, if the utility function is additive, this procedure provides tractable likelihood functions. Consider the following random specification of an additive utility function

(4.1) $$U(x;\varepsilon) = \sum_{i=1}^{m} W_i(x_i + e^{\varepsilon_i})$$

where the functions $W_i(\cdot)$, $i=1,\ldots,m$, are strictly increasing and strictly concave and $\varepsilon_1, \ldots, \varepsilon_m$ are mutually independent. Consider the general consumption pattern with demand quantities $x^* = (0,\ldots,0,x^*_{\ell+1},\ldots,x^*_m)$. The virtual prices at x^* are

(4.2) $$\xi_i = v_m \frac{\partial W_i(x_i^* + e^{\varepsilon_i})}{\partial x_i} / \frac{\partial W_m(x_m^* + e^{\varepsilon_m})}{\partial x_m} \qquad i=1,\ldots,m$$

The conditions (2.4) are

(4.3) $$\frac{\partial W_i(e^{\varepsilon_i})}{\partial x_i} \le \frac{v_i}{v_m} \frac{\partial W_m(x_m^* + e^{\varepsilon_m})}{\partial x_m} \qquad i=1,\ldots,\ell$$

and

$$(4.4)\quad \frac{\partial W_j(x_j^* + e^{\varepsilon_j})}{\partial x_j} = \frac{v_j}{v_m}\frac{\partial W_m(x_m^* + e^{\varepsilon_m})}{\partial x_m} \qquad j=\ell+1,\ldots,m-1$$

To simplify notation, denote $\Psi_i(x) = \dfrac{\partial W_i(x)}{\partial x}$, $i=1,\ldots,m$. Since $W_i(x)$ is a strictly concave function, Ψ_i must be a strictly decreasing function. Hence it follows from (4.3) and (4.4) that

$$(4.3)'\quad e^{\varepsilon_i} \geq \Psi_i^{-1}\left[\frac{v_i}{v_m}\Psi_m(x_m^* + e^{\varepsilon_m})\right] \qquad i=1,\ldots,\ell$$

and

$$(4.4)'\quad e^{\varepsilon_j} = \Psi_j^{-1}\left[\frac{v_j}{v_m}\Psi_m(x_m^* + e^{\varepsilon_m})\right] - x_j^* \qquad j=\ell+1,\ldots,m-1$$

Given x^*, the corresponding values of ε_m, such that the equation (4.4)′ holds, are determined by the inequalities

$$\Psi_j^{-1}\left[\frac{v_j}{v_m}\Psi_m(x_m^* + e^{\varepsilon_m})\right] > x_j^* \qquad j=\ell+1,\ldots,m-1$$

The feasible values of ε_m are, therefore, determined by

$$(4.5)\quad \varepsilon_m > R(x^*)$$

where

$$(4.6)\quad R(x^*) = \ln\left[\max\{0, \Psi_m^{-1}\left[\frac{v_m}{v_j}\Psi_j(x_j^*)\right] - x_m^* \quad \text{for } j=\ell+1,\ldots,m-1\}\right]$$

Since the set $\{\varepsilon_m | \varepsilon_m > R(x^*)\}$ is nonempty, the model is coherent. Denote

$$(4.7)\quad \varepsilon_j(x^*, \varepsilon_m) = \ln\left[\Psi_j^{-1}\left[\frac{v_j}{v_m}\Psi_m(x_m^* + e^{\varepsilon_m})\right] - x_j^*\right] \qquad j=\ell+1,\ldots,m-1$$

for the values of ε_m satisfying (4.5), and

$$\text{(4.8)} \quad \varepsilon_i(x^*, \varepsilon_m) = \ln \Psi_i^{-1}\left[\frac{v_i}{v_m}\Psi_m\,(x_m^* + e^{\varepsilon_m})\right] \qquad i=1,...,\ell$$

The likelihood function for this observation x^* is

$$\text{(4.9)} \quad L(x^*) = \int_{R(x^*)}^{\infty} \prod_{i=1}^{\ell} [1-F_i(\varepsilon_i(x^*, \varepsilon_m))] \cdot \prod_{i=\ell+1}^{m-1} f_i(\varepsilon_i(x^*, \varepsilon_m)) \cdot f_m(\varepsilon_m) \cdot d\varepsilon_m$$

b. Scaling Specification

The scaling procedure consists of appending multiplicative terms to the consumption vector x in the utility function. The random utility function is specified as

$$\text{(4.10)} \quad U(x;\varepsilon) = W(e^{\varepsilon_1}x_1, \ldots, e^{\varepsilon_m}x_m)$$

where the function W is strictly increasing and strictly concave on all its arguments, and the random variables $\varepsilon_1, \ldots, \varepsilon_m$ are mutually independent. Consider the general consumption pattern with demand quantities $x^* = (0,...,0, x_{\ell+1}^*, \ldots, x_m^*)$. The virtual prices at x^* are

$$\text{(4.11)} \quad \xi_i = \frac{v_m \dfrac{\partial W(0,...,0,e^{\varepsilon_{\ell+1}}x_{\ell+1}^*, \ldots, e^{\varepsilon_m}x_m^*)}{\partial x_i} e^{\varepsilon_i}}{\dfrac{\partial W(0,...,0,e^{\varepsilon_{\ell+1}}x_{\ell+1}^*, \ldots, e^{\varepsilon_m}x_m^*)}{\partial x_m} e^{\varepsilon_m}} \qquad i=1,...,m$$

where $\partial W(\bullet)/\partial x_i$ is the partial derivative of W with respect to its i^{th} argument. To simplify notation, denote $\Psi_i(\bar{x}^*, \bar{\varepsilon}) = \partial W(0,...,0, e^{\varepsilon_{\ell+1}}x_{\ell+1}^*, \ldots, e^{\varepsilon_m}x_m^*)/\partial x_i$ where $\bar{x}^* = (x_{\ell+1}^*, \ldots, x_m^*)$ and $\bar{\varepsilon} = (\varepsilon_{\ell+1}, \ldots, \varepsilon_m)$. The conditions (2.4) for the optimality of x^* are

$$\text{(4.12)} \quad \varepsilon_i \le \ln v_i - \ln v_m + \ln \Psi_m(\bar{x}^*, \bar{\varepsilon}) - \ln \Psi_i(\bar{x}^*, \bar{\varepsilon}) + \varepsilon_m \qquad i=1,...,\ell$$

and

$$(4.13)\quad \varepsilon_j = \ln v_j - \ln v_m + \ln \Psi_m(\bar{x}^*, \bar{\varepsilon}) - \ln \Psi_j(\bar{x}^*, \bar{\varepsilon}) + \varepsilon_m \qquad j=\ell+1,\ldots,m-1$$

The equations in (4.13) are functions of $x^*_{\ell+1}, \ldots, x^*_{m-1}$ and $\varepsilon_{\ell+1},\ldots,\varepsilon_m$ but not $\varepsilon_1, \ldots, \varepsilon_\ell$. The model coherency requirement will be satisfied if values of $\varepsilon_{\ell+1}, \ldots, \varepsilon_m$ exist such that the equations in (5.4) are satisfied for any given $\bar{x}^*$. A complication is that $\varepsilon_{\ell+1}, \ldots, \varepsilon_{m-1}$ may appear nonlinearly on both sides of the equations, and hence there is no guarantee that this stochastic specification may not restrict the range of demand quantities. Coherency depends on the specific functional form of the utility function W. Suppose that $\varepsilon_{\ell+1}, \ldots, \varepsilon_{m-1}$ can be solved from (4.13) as functions of $\bar{x}^*$ and ε_m with values of ε_m on some range $S(\bar{x}^*)$ which are denoted as $\varepsilon_j(\bar{x}^*, \varepsilon_m)$, $j=\ell+1, \ldots, m-1$. The likelihood function for x^* will then be

$$(4.14)\quad L(x^*) = \int_{S(\bar{x}^*)} \prod_{i=1}^{\ell} F_i(\varepsilon_i(\bar{x}^*, \varepsilon_m)) \prod_{i=\ell+1}^{m-1} f_i(\varepsilon_i(\bar{x}^*, \varepsilon_m)) |J(\bar{x}^*, \varepsilon_m)| d\varepsilon_m$$

where

$$\varepsilon_i(\bar{x}^*, \varepsilon_m) = \ln v_i - \ln v_m + \ln \Psi_m(\bar{x}^*, \bar{\varepsilon}(\bar{x}^*, \varepsilon_m)) - \ln \Psi_i(\bar{x}^*, \bar{\varepsilon}(\bar{x}^*, \varepsilon_m)) + \varepsilon_m \qquad i=1,\ldots,\ell$$

c. Indirect Utility with Scaling

It is apparent from the above discussion that both scaling and translating are useful methods to introduce consumer characteristics into demand systems derived from additive utility functions. The scaling method may be more advantageous than translating when demand equations are derived from an indirect utility (or cost) function since it does not depend specifically on the form of the indirect utility function except that the coherency conditions be satisfied. Corresponding to the utility function (4.10), the indirect utility function $V(v;\varepsilon)$ has the form[5]

$$(4.15)\quad V(v;\varepsilon) = H(v_1 e^{-\varepsilon_1}, \ldots, v_m e^{-\varepsilon_m})$$

where $H(p) = \max\{W(y) | p'y = 1\}$. The notational demand system corresponding to $V(v;\varepsilon)$, derived from Roy's identity, is

$$(4.16)\quad x_i = D_i(v_1 e^{-\varepsilon_1}, \ldots, v_m e^{-\varepsilon_m}) e^{-\varepsilon_i} \qquad i=1,\ldots,m$$

The dual approach specifies either the functional form of the indirect function H or the demand functions D_i, $i=1,\ldots,m$. The stochastic elements will then be incorporated multiplicatively in the functions (4.15) and (4.16). The derivation of the likelihood function proceeds as follows. Consider the consumption pattern with $x^* = (0,\ldots,0,x^*_{\ell+1}, \ldots, x^*_m)$ where $x^*_j > 0$, $j=\ell+1,\ldots,m$. The virtual prices ξ_i, $i=1,\ldots,\ell$ for the first ℓ goods at x^* are characterized by the following relations:

$$(4.17)\quad 0 = D_i(\xi_1 e^{-\varepsilon_1}, \ldots, \xi_\ell e^{-\varepsilon_\ell}, v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m}) e^{-\varepsilon_i} \qquad i=1,\ldots,\ell$$

$$(4.18)\quad x^*_j = D_j(\xi_1 e^{-\varepsilon_1}, \ldots, \xi_\ell e^{-\varepsilon_\ell}, v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m}) e^{-\varepsilon_j} \qquad j=\ell+1,\ldots,m$$

The factors $e^{-\varepsilon_i}$ in (4.17) can be dropped. Firstly, solve the factors $\xi_i e^{-\varepsilon_i}$, $i=1,\ldots,\ell$ from (4.17) as functions of $(v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m})$,

$$(4.19)\quad \xi_i e^{-\varepsilon_i} = h_i(v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m}) \qquad i=1,\ldots,\ell$$

Substituting (4.19) into (4.18), they imply

$$(4.20)\quad x^*_j = D_j(h_1, \ldots, h_\ell, v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m}) e^{-\varepsilon_j}, \qquad j=\ell+1,\ldots,m-1$$

which involves only the random variables $\varepsilon_{\ell+1}, \ldots, \varepsilon_m$. The regime conditions $\xi_i \le v_i$, $i=1,\ldots,\ell$ become

$$(4.21)\quad \varepsilon_i \le \ln v_i - \ln h_i(v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m}), \quad i=1,\ldots,\ell$$

Suppose that the model is coherent and $\varepsilon_{\ell+1}, \ldots, \varepsilon_{m-1}$ can be solved from (4.20) as functions of $x^*_1, \ldots, x^*_{m-1}$ and ε_m where ε_m has a range of effective values $S(x^*)$. Denote these functions as

$\varepsilon_i(x^*, \varepsilon_m)$, $i = \ell+1, \ldots, m-1$. Denote also $\varepsilon_i(x^*, \varepsilon_m) = \ln v_i - \ln h_i(v_{\ell+1} e^{-\varepsilon_{\ell+1}}, \ldots, v_m e^{-\varepsilon_m})$ for $i = 1, \ldots, \ell$. The likelihood function for x^* will be

$$(4.22)\quad L(x^*) = \int_{S(x^*)} \prod_{i=1}^{\ell} F_i(\varepsilon_i(x^*, \varepsilon_m)) \prod_{i=\ell+1}^{m-1} f_i(\varepsilon_i(x^*, \varepsilon_m)) |J(x^*, \varepsilon_m)| f_m(\varepsilon_m)\, d\varepsilon_m$$

In practice, the difficulty of this approach lies in the derivation of the virtual price functions (4.19) and the equations $\varepsilon_i(x^*, \varepsilon_m)$ from (4.20). The primal approach is somewhat simpler in that the derivation of the virtual prices is straightforward for any specified direct utility function. In a preliminary paper, Lee [9] has shown that the linear expenditure and the translog indirect utility functions with the scaling specification provide computationally tractable likelihood functions. This approach is also useful in production analysis (see Lee [9]) and can be extended for the analysis of discrete choice models; see Haneman [7] and Lee and Chiang [10].

5. Estimation of the Linear Expenditure System: The Demand for Food in Indonesia

a. Stochastic Specification of the LES System

To demonstrate this approach to estimating demand systems with many binding non-negativity constraints, a linear expenditure system of food demand equations is estimated using a household expenditure survey from Indonesia. The linear expenditure system estimated is derived from a Klein-Rubin-Stone-Geary direct utility function of the form

$$(5.1) \quad u(x) = \sum_{i=1}^{m} \alpha_i \ln(x_i - \beta_i)$$

where $\alpha_i > 0$. The implied notional expenditure share equations are

$$(5.2) \quad v_i x_i = v_i \beta_i + \theta_i (1 - \sum_{j=1}^{m} v_j \beta_j) \qquad i=1,\ldots,m$$

where $\theta_i = \alpha_i / \sum_{j=1}^{m} \alpha_j$. Corner solutions (zero demands) can occur only if some of the parameters β_i are negative. Goods for which the corresponding β_i is negative are referred to as "inessential" because of the common interpretation of the β's as representing subsistence or committed expenditure.

Variation in tastes across consumers is introduced into the utility function (5.1) by treating the α_i parameters as stochastic

$$(5.3) \quad \alpha_i = e^{\varepsilon_i}$$

where the disturbances ε_i are mutually independent and are normally distributed $N(\gamma_i, \sigma_i^2)$, $i=1,\ldots,m$. Demographic variables are introduced into the demand system by treating γ_i, the means of ε_i in (5.3), as linear functions of demographic variables. Since the additive form of the utility function in (5.1) is invariant with respect to linear transformation, the normalization $\gamma_m = 0$ is made. With this normalization, the system is observationally equivalent to the specification that $\alpha_m = 1$ and $\varepsilon_1, \ldots, \varepsilon_{m-1}$ are jointly normally distributed with an error component structure, i.e., the covariances of ε_i and ε_j are the same for all $i, j=1,\ldots,m-1$ and $i \neq j$. At the sample observation $x^* = (0,\ldots,0,x^*_{\ell+1}, \ldots, x^*_m)$ with $x_i^* > 0$, $i=\ell+1,\ldots,m$, the virtual prices are

$$(5.4) \quad \xi_i = v_i \frac{s_m^* - v_m \beta_m}{s_i^* - v_i \beta_i} e^{\varepsilon_i - \varepsilon_m} \qquad i=1,\ldots,m-1$$

where $s_i^* = v_i x_i^*$ is the expenditure share of the i^{th} commodity. The optimality conditions for the x^* are, after logarithmic transformation,

$$(5.5) \quad \varepsilon_i \leq \ln(-v_i \beta_i) - \ln(s_m^* - v_m \beta_m) + \varepsilon_m \qquad i=1,\ldots,\ell$$

and

$$(5.6) \quad \varepsilon_j = \ln(s_j^* - v_j \beta_j) - \ln(s_m^* - v_m \beta_m) + \varepsilon_m \qquad j=\ell+1,\ldots,m-1$$

Since $s_m^* = 1 - \sum_{j=\ell+1}^{m-1} s_j^*$, the Jacobian of the transformation (5.6) from $\varepsilon_{\ell+1}, \ldots, \varepsilon_{m-1}$ to $s^*_{\ell+1}, \ldots, s^*_{m-1}$ is $\sum_{j=\ell+1}^{m} (s_j^* - v_j \beta_j) / \prod_{j=\ell+1}^{m} (s_j^* - v_j \beta_j)$. The likelihood function for x^* is

$$(5.7) \quad L(x_i^*)$$

$$= \frac{\sum_{j=\ell+1}^{m} (s_j^* - v_j \beta_j)}{\prod_{j=\ell+1}^{m} (s_j^* - v_j \beta_j)} \int_{-\infty}^{\infty} \prod_{i=1}^{\ell} \Phi\left[\frac{\varepsilon_i(x^*, \varepsilon_m) - \gamma_i}{\sigma_i}\right] \cdot \prod_{j=\ell+1}^{m-1} \frac{1}{\sigma_j} \phi\left[\frac{\varepsilon_j(x^*, \varepsilon_m) - \gamma_j}{\sigma_j}\right] \frac{1}{\sigma_m} \phi\left[\frac{\varepsilon_m}{\sigma_m}\right] d\varepsilon_m$$

where $\varepsilon_i(x^*, \varepsilon_m) = \ln(s_i^* - v_i \beta_i) - \ln(s_m^* - v_m \beta_m) + \varepsilon_m$, $i=1,\ldots,m-1$; Φ is the standard normal distribution function and ϕ is the standard normal density function. With the normality assumption, the integral can be numerically evaluated with the Gaussian quadrature formula; see Stroud and Secrest [16].

b. Data and Results

A sample of 1150 households were drawn from the 1978 Socioeconomic Survey of Indonesia (SUSENAS), a national probability sample of households. Food consumption (purchased and home-produced) of nearly 100 separate items in the seven days prior to the date of enumeration is aggregated into seven categories: tubers, fruits, animal products (meat and dairy), fish, vegetables, grains and others. A village was assumed to represent a distinct market, and the average price of every disaggregate item is calculated as the average price of the commodity consumed by the sampled households in the village. Price indices are computed by geometrically weighting component prices with the average budget shares of a larger administrative area, the kabupaten (regency)[6]. There are 300 kabupatens in the sample. The absence of data on most non-food prices means that we must impose the assumption that foods and non-foods are separable in the utility function. Three demographic variables are identified--the number of household members 4 years of age and under (infants), the number aged 5 through 14 (children), and those of age 15 and above (adults). Table 1 provides summary statistics on food consumption shares and normalized (by total food expenditure) prices, as well as demographic variables.

As Table 1 indicates, six of the seven foods were not consumed by at least one household during the reference period. Half of the sampled households did not consume animal products, and one third did not consume tubers or fruit. Only grain was consumed by all households. As noted above, it is required that all goods which have zero demands have a corresponding β_i

which is negative in order that the utility function (5.1) be defined. It follows that only grains can be an "essential" good.

Table 2 provides maximum likelihood estimates of the utility function with demographic effects.[7] The coefficients β_i are indeed significantly negative for all goods except grain and all the asymptotic t-ratios are quite large. The coefficients on the adult and children demographic variables are all statistically significant. The infant variable is significant only in the case of animal products and fish.[8] Parameter estimates are interpreted below in terms of elasticities.

Table 3 presents estimates of the effects of incremental household members (by type) on consumption. Adding an infant to a household having mean demographic characteristics (see Table 1) reduces the consumption of all goods except grain and others. There is a particularly sharp fall in the consumption of animal products (12.6 percent) and fish (8.00 percent), both important but expensive sources of protein. Adding a child rather than an infant leads to an even greater reduction in animal product consumption (20.9 percent). Fruit consumption is now sharply reduced and only grain consumption rises. Only for tubers, fruit, animal products and fish is consumption response monotonic in response to an additional household member in age order: infant, child, adult. Tubers, fruit and animal products consumption response rises in magnitude with higher aged incremental household members while fish response falls in magnitude. Only grain, which has the largest average expenditure, is consumed in even greater amount with incremental household members of any type.

Table 4 provides the matrix of compensated and uncompensated price elasticities as well as income elasticities.[9] These elasticities were calculated for a representative household having sample mean demographic characteristics, random terms and shares, and virtual prices which support those shares given sample mean demographic characteristics. Interpretation of these elasticities requires one to recognize the properties of the LES functional form. As is well known, the non-negativity of the α_i's in (5.1) rules out inferiority, and concavity requires that every good must be a substitute for every other good. Furthermore, the additivity of preferences implies--through Pigou's Law (Deaton [5])--that for "large" numbers of goods, income and price elasticities are approximately proportional. With this in mind, note that price and expenditures elasticities are highest (in absolute value) for animal products and lowest for grain. Grain is the only price inelastic good in the set of foods and all other goods are price elastic. Animal products are commonly found to be highly income elastic in developing countries. Among these seven foods, tubers, vegetables, grain and others are all "necessities" as the expenditure elasticities are less than one. Fruit, animal products and fish are all "luxury" goods. All the elasticities seem sensible, but one might wonder whether a non-additive model might result in very different elasticity measurements. We are currently investigating non-additive models, but they are much more complicated to estimate.

6. Summary

In this paper we have presented **coherent** stochastic specifications for direct and indirect utility functions which result in computationally tractable demand systems subject to binding non-negativity constraints. The implied likelihood functions do not involve multiple integrals. As an example of one of the specifications, a seven-commodity demand system for food was estimated using household consumption data from Indonesia. These approaches promise to greatly relieve the computational burden of estimating large demand systems subject to binding non-negativity constraints that were limitations in our earlier work (Lee and Pitt [11]) and that of Wales and Woodland [18]. Although the example estimated here was for the relatively restrictive linear expenditure system, less restrictive functional forms, while computationally more burdensome than the LES, may still be suitable for applied demand analysis.

Table 1

Summary Statistics

	Mean	Standard Deviation	Frequency of Zero Consumption
Tubers Share	.03126	0.5205	473
Fruit Share	.03638	.04646	419
Animal Products Share	.05591	.08720	576
Fish Share	.11112	.08937	112
Vegetables Share	.12853	.07023	14
Others Share	.18061	.07921	1
Grain Share	.45619	.16298	0
Tuber Price	1.27625	1.05631	
Fruit Price	1.18679	1.03619	
Animal Products Price	1.0964	.84715	
Fish Price	1.14076	.89651	
Vegetable Price	1.14780	.77964	
Others Price	1.21362	1.02743	
Grain Price	1.12126	.80051	
Infants[a]	.77565	.85407	
Children[a]	1.60870	1.39685	
Adults[a]	3.0330	1.30632	

Sample size = 1150

[a] Infants, children and adults are household members aged 0 to 4, 5 to 15, and 16 years of age and above, respectively.

Table 2

Parameter Estimates

	Tubers	Fruit	Animal Products	Fish	Vegetables	Others	Grain
β	-0.0267 (-8.8708)	-0.0474 (-11.8979)	-.2343 (-11.9465)	-.0536 (-12.2722)	-.0327 (-11.0840)	-.0299 (-9.4441)	.0122 (3.3862)
Constant	-1.9602 (-13.4296)	-1.3167 (-12.3139)	-.1338 (-1.2209)	-.6223 (-7.1550)	-.6226 (-9.2591)	-.4649 (-7.3691)	--
Infants	-.0407 (-.9398)	-.0361 (-1.0944)	-.0623 (-2.2052)	-.0890 (-2.9782)	-.0395 (-1.7160)	-.0057 (-.2600)	--
Children	-.0953 (-3.6856)	-.1167 (-5.6813)	-.1159 (-6.5021)	-.0904 (-5.1844)	-.0911 (-6.4036)	-.0694 (-5.1035)	--
Adults	-.1104 (-3.7798)	-.1252 (-5.7829)	-.1128 (-6.3366)	-.0731 (-3.6181)	-.0722 (-4.6829)	-.0448 (-2.9864)	--
σ	1.0286 (19.3572)	.7309 (25.5662)	.5546 (21.8458)	.6556 (25.6824)	.4391 (25.9058)	.3702 (26.9344)	.6421 (33.8889)
$\hat{\theta}$ (at mean z)	.0287	.0506	.1683	.1189	.1237	.1673	.3425
$\bar{\theta}$ (at mean z)	.0398	.0574	.1666	.1271	.1198	.1600	.3292

Note:

(1) 't' statistics are in brackets

(2) $\hat{\theta}_i = e^{\bar{z}\gamma_i} / \sum_{j=1}^{m} e^{\bar{z}\gamma_j}$.

(3) $\bar{\theta}_i = E_\varepsilon(e^{\bar{z}\gamma_i + \varepsilon_i} / \sum_{j=1}^{m} e^{\bar{z}\gamma_j + \varepsilon_j})$.

Table 3

Demographic Effects

Percent Change in Consumption of [a]:

In Response to Incremental:	Tubers	Fruit	Animal Products	Fish	Vegetables	Other	Grain
Infant	-1.33	-1.19	-12.57	-8.00	-1.25	+2.87	+2.89
	(-0.04)	(-0.04)	(-0.70)	(-0.89)	(-0.16)	(+0.52)	(+1.32)
Child	-4.03	-9.79	-20.90	-4.05	-3.73	-1.04	+6.06
	(-0.13)	(-0.36)	(-1.17)	(-0.45)	(-0.48)	(-0.19)	(+2.77)
Adult	-6.66	-12.59	-22.73	-2.80	-2.41	+0.94	+5.24
	(-0.21)	(-0.46)	(-1.27)	(-0.31)	(-0.31)	(+0.17)	(+2.39)

(a) Percent change relative to sample mean shares. Absolute change in expenditure share are in parentheses underneath percent change in consumption.

Table 4

Uncompensated Price Elasticities

Quantities/Prices	Tubers	Fruit	Animal Products	Fish	Vegetables	Other	Grain
Tubers	-1.1802	.0266	.1482	.0389	.0286	.0326	-.0125
Fruit	.0080	-1.7570	.2246	.0590	.0434	.0494	-.0190
Animal Products	.0174	.0873	-3.4022	.1277	.0940	.1069	-.0411
Fish	.0062	.0310	.1727	-1.3365	.0334	.0380	-.0146
Vegetables	.0055	.0279	.1553	.0408	-1.2130	.0341	-.0131
Others	.0053	.0266	.1495	.0393	0.0289	-1.1637	-.01265
Grain	.0043	.0217	.1212	.0318	.0234	.0266	-.9603

Compensated Price Elasticities

Quantities/Prices	Tubers	Fruit	Animal Products	Fish	Vegetables	Other	Grain
Tubers	-1.1515	.0600	.1995	.1409	.1466	.1983	.4061
Fruit	.0515	-1.7064	.3024	.2136	.2223	.3007	.6157
Animal Products	.1115	.1968	-3.2340	.4622	.4809	.6505	1.3319
Fish	.0396	.0699	.2325	-1.2176	.1709	.2312	.4733
Vegetables	.0356	.0629	.2091	.1477	-1.0893	.2079	.4258
Others	.0343	.0605	.2013	.1422	.1480	-.9964	.4099
Grain	.0278	.0491	.1632	.1153	.1199	.1622	-.6377

Total Expenditure Elasticities

Tubers	Fruit	Animal Products	Fish	Vegetables	Other	Grain
0.9177	1.3913	3.099	1.070	.9623	.9263	.7509

Footnotes

(*) The first author appreciates having financial support from the National Science Foundation under grant number SES-8510473 for his research. Both authors acknowledge computational support provided by an NSF supercomputer grant. We also thank the Biro Pusat Statistik in Jakarta for access to data.

1 Household purchase (as opposed to consumption) data may also be characterized by many zeroes over a reference period. Under certain conditions, our model estimated with purchase data may be interpretable as a short-term purchase model. The estimation results we report in this paper use consumption data.

2 Before Wales and Woodland, this stochastic specification had been studied by Theil and Neudecker [17] and Barten [3].

3 Alternative distributions may also be feasible as long as their support is not bounded from above (see (3.5)).

4 The normalization $\sum_{i=1}^{m} \varepsilon_i = 0$ used by Wales and Woodland implies that the disturbances ε_i must be correlated and their covariance matrix is singular. Such a specification does not result in a computationally tractable likelihood function.

5 Barten [2] previously obtained this reuslt.

6 If a commodity's price was unavailable for a village, it was taken to be the average kabupaten price.

7 Numerical optimization was accompanied with both the Powell and Berndt, Hall, Hall and Hausman algorithms in the Goldfeld and Quandt optimization package GQOPT. The computation was carried out on both an IBM PC-AT and a Cray-1 supercomputer at the University of Minnesota.

8 We experimented with the model by treating demand as continuous dependent variables, thus neglecting their discrete nature. Using the MLE of Table 2 as initial values, parameter estimates changed sharply after only a few iterations. This suggests that explicitly treating the binding non-negativity constriants importantly affects the estimates.

9 The own price elasticities for the LES system are $\frac{\partial \ln x_i}{\partial \ln v_i} = -1 + \frac{(1-\theta_i)\beta_i}{x_i}$, $i=1,\ldots,m$. The cross price elasticities are $\frac{\partial \ln x_i}{\partial \ln v_j} = -\theta_i \frac{v_j \beta_j}{v_i x_i}$, $i \neq j$, $i,j=1,\ldots,m$, and the income elasticities are $\frac{\partial \ln x_i}{\partial \ln M} = \frac{\theta_i}{v_i x_i}$, $i=1,\ldots,m$.

REFERENCES

[1]

Amemiya, T.
Multivariate Regression and Simultaneous Equation Models When the Dependent Variables Are Truncated Normal.
Econometrica, 42:999-1012, 1974.

[2]

Barten, A. P.
Family Composition, Price and Expenditure Patterns,
in *Econometric Analysis for National Economic Planning*, P. E. Hart, G. Mills and J. K. Whitaker, editors.
Butterworth, London, 1964.

[3]

Barten, A. P.
Estimating Demand Equations.
Econometrica, 36:213-251, 1968.

[4]

Burtless, G. and J. A. Hausman.
The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income-Tax Experiment.
Journal of Political Economy, 86:1103-1130, 1978.

[5]

Deaton, A. S.
A Reconsideration of the Empirical Implications of Additive Preferences.
Economic Journal, 84:338-348, 1974.

[6]

Gourieroux, C., J. J. Laffont and A. Monfort.
Coherency Conditions in Simultaneous Linear Equations Models with Endogenous Switching Regimes.
Econometrica, 48:675-695, 1980.

[7]

Hanemann, W. M.
Discrete/Continuous Models of Consumer Demand.
Econometrica, 52:541-561, 1984.

[8]
Lee, L. F. and M. M. Pitt.
Microeconometric Models of Consumer and Producer Demand with Limited Dependent Variables.
Economic Development Center, University of Minnesota, 1984.

[9]
Lee, L. F.
Microeconometric Models of Consumer and Production Demand With Binding Non-Negativity Constraints: The Disaggregated Commodities Cases.
manuscript, Department of Economics, University of Minnesota, 1985.

[10]
Lee, L. F. and J. W. Chiang.
Discrete/Continuous Choice and Purchase Decision Econometric Models for Consumer Demand.
Department of Economics, University of Minnesota, 1986.

[11]
Lee, L. F. and M. M. Pitt.
Microeconometric Demand Systems With Binding Non-Negativity Constraints: The Dual Approach.
Econometrica, forthcoming:September issue, 1986.

[12]
McFadden, D.
Conditional Logit Analysis of Qualitative Choice Behavior,
in *Frontier in Econometrics*P. Zarembka, editor.
Academic Press, New York, 1973.

[13]
Pollak, R. A. and T. J. Wales.
Estimation of the Linear Expenditure System.
Econometrica, 37:611-628, 1969.

[14]
Pollak, R. A. and T. J. Wales.
Demographic Variables in Demand Analysis.
Econometrica, 49:1533-1551, 1981.

[15]
Stone, J. R. N.
Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand.
Economic Journal, 64:511-527, 1954.

[16]
Stroud, A. H. and D. Secrest.
Gaussian Quadrature Formulas.
Prentice Hall, Englewood Cliffs, N.J., 1966.

[17]
Theil, H. and H. Neudecker.
Substitution, Complementarity and Residual Variation around Engel Curves.
The Review of Economic Studies, 25:114-123, 1957.

[18]
Wales, T. J. and A. D. Woodland.
Estimation of Consumer Demand Systems With Binding Non-Negativity Constraints.
Journal of Econometrics, 21:263-285, 1983.