DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURAL AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

Working Paper No. 911

THE U.S. INCOME DISTRIBUTION AND GORMAN ENGEL CURVES FOR FOOD

J. T. LaFrance
University of California, Berkeley

G. K. Agnew
University of Arizona

T. K. M. Beatty
University of California, Berkeley

R. D. Pope
Brigham Young University

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
2001

The U.S. Income Distribution and Gorman Engel Curves for Food

J. T. LaFrance
University of California, Berkeley

G. K. Agnew
University of Arizona

T. K. M. Beatty
University of California, Berkeley

R. D. Pope
Brigham Young University

Abstract

A method for nesting, estimating and testing for the rank and functional form of the income terms in an incomplete system of aggregable and integrable demand equations is derived. The Maximum entropy (MAXENT) procedure is applied to the problem of inferring the U.S. income distribution using annual time series data on quintile and top five percentile income ranges and intra-quintile and top five percentile mean incomes. The MAXENT results are compared with those obtained from a parametric method utilizing a truncated three-parameter lognormal distribution. The two estimates for the year-to-year income distribution are combined with annual time series data on the U.S. consumption of and retail prices for twenty-one food items to estimate the rank and functional form of the income terms in U.S. food demand over the period 1919-1995, excluding 1942-1946 to allow for the structural impacts of World War II.

Corresponding Author:

Professor Jeffrey T. LaFrance
Department of Agricultural and Resource Economics
207 Giannini Hall / MC 3310
University of California
Berkeley, CA 94720-3310
Voice: 510-643-5416
Fax: 510-643-8911
E-mail: jeff_lafrance@are.berkeley.edu

The U.S. Income Distribution and Gorman Engel Curves for Food

1. Introduction

Following Muellbauer's (1975) extension of the Gorman polar form to a nonlinear function of income to obtain the price independent generalized linear (PIGL) and price independent generalized logarithmic (PIGLOG) functional forms, much progress has been made in the past 25 years on aggregation theory in consumption. The Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980) implements Muellbauer's results to produce demands with budget shares expressed as functions of linear and quadratic terms in the logarithm of prices and a linear term in the logarithm of income. The AIDS and its linear approximation (LA-AIDS) have been linchpins in applied demand analysis since their introduction. Most applications of the AIDS and LA-AIDS either assume separability and estimate a complete system of demands for a disaggregate group of commodities as functions of prices for the goods in the group and total expenditure on the group, or estimate a complete system of demands with highly aggregated commodities as functions of aggregate price indices and total consumption expenditures (hereafter, income, which we denote by *m*).

Shortly after the article by Deaton and Muellbauer, in a remarkable and elegant contribution to the festschrift to Sir Richard Stone, Gorman (1981) derived the set of functional forms for demand models that can be written in terms of any additive set of functions of income. Any complete system of demand equations in the class of "Gorman Engel curves" must satisfy two properties in addition to homogeneity, adding up and symmetry. First, if the number of independent functions of income is at least three, then the functions all must be either (a) polynomials in income, (b) polynomials in some non-integer power of income, (c) polynomials in the natural logarithm of income, or (d) a series of sine and cosine functions of the natural logarithm of income. Second, the number of "linearly independent" functions of income in this class of demand systems at most equals three, where linear independence refers to the rank of the matrix of price functions that premultiply the income functions. One important implication is that theoretically consistent demand aggregation in models that have full column rank for this matrix requires three summary statistics from the distribution of income to estimate the demand parameters with aggregate data.

Gorman (1981) also conjectured that second-order polynomials are the most general non-degenerate cases of demand systems that have full rank three. Pursuing this conjecture by exploiting the methods of van Daal and Merkies (1989), Lewbell (1990) was able to show that all full rank three generalizations of Muellbauer's PIGL and PIGLOG demand models are quadratic forms analogous to the quadratic expenditure system (QES) developed by Howe, Pollak and Wales (1979) and perfected by van Daal and Merkies (1989). Lewbell (1990) also derived a full rank three trigonometric model.

All of the above results on the rank of the coefficient matrix and the functional form of the income terms in the class of Gorman Engel curve demand models require the adding up property of a complete demand system. However, often we are interested in the demands for a subset of goods that make up only part of the consumption budget. In such a case, separability is a strong assumption, and it is undesirable to impose strong restrictions without good reason or prior evidence. Without separability, there is little reason to impose the same functional form on the demand equations for the goods of interest and all of the other goods for which we have little or no price or quantity information. This implies that the above results cannot be applied directly to incomplete demand systems.

In an ambitious paper, Gorman (1965; 1995) considered the structure of the demands for groups of goods in which each group's total expenditure is a function of income and a set of aggregate price indices for each group, and derived the restrictions on the individual demand equations and the properties of the indirect utility function under this set of restrictions. Independently and more recently, but along a similar line of thought, Epstein (1982), LaFrance (1985) and LaFrance and Hanemann (1989) developed a theory for the *weak integrability* of the demands for a single proper subset of all goods that does not exhaust the consumer's budget, regardless of the number of prices that enter the demand equations. The conditions for weak integrability of an incomplete demand system are that the demands are positive valued, 0° homogeneous in all prices and income, the budget restriction takes the form of a strict inequality (not all of income is exhausted by the subset of goods under study), and the submatrix of Slutsky substitution terms associated with this subset of demands is symmetric and negative semidefinite. These conditions exhaust the properties implied by consumer theory for any proper subset of all goods and are necessary and sufficient for the recovery of the conditional preference functions (both direct and indirect) for those goods, with prices of all other goods acting as conditioning variables (LaFrance (1985); LaFrance and Hanemann (1989)). *Inter alia*, the set of incomplete demand models that satisfy weak integrability is much richer than the corresponding set of integrable complete demand systems.

This paper exploits the richness of the set of weakly integrable demand models to extend aggregation in nonlinear functions of income to incomplete demand systems for the PIGL and PIGLOG members of Gorman Engel curves. These extensions permit us to develop a method to nest weakly integrable LA-AIDS, AIDS, quadratic AIDS (QAIDS), quadratic PIGL (QPIGL), and extended QES[1] models to simultaneously test for and estimate both the rank and functional form of the income terms in aggregable incomplete demand systems.

As noted above, a full rank three Gorman Engel curve demand model requires three summary statistics from the income distribution, e.g., for a QPIGL model in expenditure form we need the cross-sectional means of $m_h^{1-\kappa}$, m_h, and $m_h^{1+\kappa}$, where m_h is the income level of family h, $h = 1, ..., H$, say, and κ is the PIGL coefficient on income, while for a QAIDS model we need the means of m_h, $m_h \ln(m_h)$, and $m_h[\ln(m_h)]^2$. To calculate

[1] "Extended QES" indicates that *supernumerary income* is income minus a quadratic form in prices and that there is an $n{\times}n$ matrix of price effects in addition to the intercepts in the QES demands.

these means, we need information on the distribution of income. The U.S. Bureau of the Census annually publishes the quintile ranges, intra-quintile means, top five-percentile lower bound for income, and the mean income within the top five-percentile range for all U.S. families. We use the Maximum entropy (MAXENT) procedure to obtain annual information theoretic density functions that satisfy each of these percentile and conditional mean conditions for the period 1910-1999. The MAXENT densities and the resulting food demand estimates are compared with those obtained from a parametric truncated three-parameter lognormal distribution for each year.

The income distribution estimates are combined with aggregate annual time series data on per family U.S. food expenditures for 21 individual food items over the period 1919-1995, excluding 1942-1946 to account for the structural impacts of World War II.[2] In addition to annual measures of food expenditures, prices, and the income distribution, we incorporate measures for the distribution of the U.S. population by age and the ethnicity of the U.S. population in the incomplete demand model's specification. The results of the empirical application strongly suggest that a full rank three model is essential, and that the QAIDS is strongly rejected in favor of an extended QES.

The rest of the paper is organized as follows. The next section extends the aggregation results of Gorman and others to incomplete demand systems that can be written in a PIGL/PIGLOG form. The third section describes the implementation of the MAXENT procedure estimating the U.S. income distribution and discusses estimation of the parametric truncated three-parameter lognormal distribution. Section 4 presents a summary and discussion of a subset of the empirical results, focising primarily on the

---

[2] See LaFrance (1999a, 1999b) for empirical evidence for the exclusion of World War II and the stability of U.S. food demands over this long sample period. The twenty-one food items included in the data set can be conveniently grouped into four categories: (1) *dairy products*, including fresh milk and cream, butter, cheese, ice cream and frozen yogurt, and canned and dried milk; (2) *meats, fish and poultry*, including beef and veal, pork, other red meat, fish, and poultry; (3) *fruits and vegetables*, including fresh citrus fruit, fresh noncitrus fruit, fresh vegetables, potatoes and sweet potatoes, processed fruit, and processed vegetables; and (4) *miscellaneous foods*, including fats and oils excluding butter, eggs, cereals, sugar and sweeteners, and coffee, tea and cocoa.

rank of the demand model and the functional form of the income terms. The final section summarizes the findings of the paper and discusses possible limitations of the analysis and possible directions for further research. Proofs of our main lemmas are contained in an Appendix.

2. Nesting LA-AIDS, AIDS and QAIDS within a QPIGL-IDS

In the two decades since its introduction by Deaton and Muellbauer, the AIDS has been widely used in demand analysis. The vast majority of empirical applications follows Deaton and Muellbauer's suggestion and replaces the translog price index that deflates income with Stone's index, which generates the LA-AIDS. Although Deaton and Muellbauer (1980: 317-320) cautioned against and avoided the practice, most empirical applications of the LA-AIDS include tests for and the imposition of an approximate version of Slutsky symmetry by restricting the matrix of logarithmic price coefficients to be symmetric. Important examples include Anderson and Blundell (1983), Buse (1998), Moschini (1995), Moschini and Meilke (1989), and Pashardes (1993).[3] In this section, we derive the conditions for integrability of LA-AIDS and a simple method for nesting the homothetic integrable solution within a class of homothetic PIGL demand models. We then extend this nesting procedure to non-homothetic PIGL and QPIGL forms.

Let $\boldsymbol{p}$ be the n-vector of market prices for goods, let u be the utility index, let $e(\boldsymbol{p},u)$ be the consumer's expenditure function, and let $\boldsymbol{w}$ be the n-vector of budget shares. If it is integrable, then the LA-AIDS can be written in matrix notation as

$$(1) \qquad \boldsymbol{w} = \frac{\partial \ln e(\boldsymbol{p},u)}{\partial \ln \boldsymbol{p}} = \alpha + \boldsymbol{B} \ln \boldsymbol{p} + \gamma \left[\ln e(\boldsymbol{p},u) - (\ln \boldsymbol{p})' \frac{\partial \ln e(\boldsymbol{p},u)}{\partial \ln \boldsymbol{p}} \right]$$

[3] However, see Browning and Meghir (1991) for an application of estimating the integrable AIDS, using the LA-AIDS with a symmetric matrix of log-price coefficients to obtain starting values for the nonlinear estimation procedure.

where α and γ are n-vectors and $\boldsymbol{B}$ is an $n\times n$ matrix of parameters. At various points in the paper, it proves to be helpful to change variables from quantities, prices, expenditures, budget shares, and income to particular transformations of these variables. In the present situation, it is most useful to define $\boldsymbol{x} \equiv \ln(\boldsymbol{p})$ and $y(\boldsymbol{x},u) \equiv \ln[e(\boldsymbol{p}(\boldsymbol{x}),u)]$, where $\boldsymbol{p}(\boldsymbol{x}) \equiv [e^{x_1} \cdots e^{x_n}]'$. With these definitions, we can rewrite (1) in the form

(2) $$(\boldsymbol{I} + \gamma\boldsymbol{x}')\frac{\partial y(\boldsymbol{x},u)}{\partial \boldsymbol{x}} = \alpha + \boldsymbol{B}\boldsymbol{x} + \gamma y(\boldsymbol{x},u) .$$

This observation leads to our first result, which identifies conditions for the (local) integrability of the LA-AIDS.[4]

Lemma 1. *If the LA-AIDS is integrable over an open set* $\theta \subset \mathbb{R}^n$ *with a nonempty interior and such that* $1+\boldsymbol{\gamma}'\boldsymbol{x} \neq 0\ \forall\ \boldsymbol{x} \in \theta$, *then either (a)* $\gamma \neq \boldsymbol{0}$ *and* $\boldsymbol{B} = \beta_0\gamma\gamma'$ *for some* $\beta_0 \in \mathbb{R}$, *or (b)* $\gamma = \boldsymbol{0}$ *and* $\boldsymbol{B} = \boldsymbol{B}'$. *In case (a), the logarithmic expenditure function has the form*

$$y(\boldsymbol{x},u) = \alpha'\boldsymbol{x} + \beta_0\left[(1+\gamma'\boldsymbol{x})\ln(1+\gamma'\boldsymbol{x}) - \frac{\gamma'\boldsymbol{x}}{(1+\gamma'\boldsymbol{x})}\right] + (1+\gamma'\boldsymbol{x})u ,$$

while in case (b) it has the form,

$$y(\boldsymbol{x},u) = \alpha'\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x} + u .$$[5]

---

[4] Although this lemma only requires conditions that are satisfied locally, it is shown in the Appendix that the set θ covers all of n-dimensional space except for an (n-1)-dimensional hyperplane (which has Lesbesgue measure zero in n-space).

[5] Case (a), where the log-income coefficients do not vanish, but the log-price coefficient matrix has rank one, is too restrictive to be of empirical interest. However, this case reveals an interesting structural property. In particular, it is characterized by a system of linear identities among *budget shares*,

(3) $$\boldsymbol{w} \equiv \alpha + \gamma(w_1 - \alpha_1)/\gamma_1\ \forall\ (\boldsymbol{p},m),$$

where, without loss in generality, we have assumed that $\gamma_1 \neq 0$. Recall that the linear expenditure system (LES) is characterized by a system of linear identities among *expenditures*, say,

(4) $$\boldsymbol{e} \equiv \alpha + \gamma(e_1 - \alpha_1)/\gamma_1\ \forall\ (\boldsymbol{p},m),$$

Case (b), which produces a homothetic demand model, is the solution of interest in this paper. In particular, this solution has exactly the same structure as the homothetic LIDS in LaFrance (1985). If one is willing to forgo symmetric functional forms for *all* demands, which is a relatively minor consideration in the case of estimating the demands for a proper subset of all goods, this suggests a simple way to nest the homothetic LA-AIDS and LIDS with Box-Cox transformations in an IDS framework. To develop this method, suppose that the model applies to n out of $N \geq n+1$ goods and define $m(\kappa) \equiv (m^{\kappa} - 1)/\kappa$, $p_i(\lambda) \equiv (p_i^{\lambda} - 1)/\lambda$, and $\boldsymbol{p}(\lambda) \equiv [p_1(\lambda) \cdots p_n(\lambda)]'$. Assume that m and $\boldsymbol{p}$ are deflated, with a common deflator that is a known, positive valued and 1° homogeneous function of (at least some of) the prices of all other goods, say, $\pi(\tilde{\boldsymbol{p}})$. Under these conditions, we can write a class of weakly integrable, homothetic PIGL-IDS models in budget share form as

$$(6) \qquad \boldsymbol{w} = m^{-\kappa} \boldsymbol{P}^{\lambda}[\alpha + \boldsymbol{B}\boldsymbol{p}(\lambda)],$$

where $\boldsymbol{P}^{\lambda} \equiv \text{diag}[p_i^{\lambda}]$ is a diagonal matrix with typical diagonal element p_i^{λ}. Using the integration techniques detailed in LaFrance and Hanemann (1989), it can be shown that the expenditure function for this PIGL-IDS satisfies

$$(7) \qquad \mathrm{e}(\boldsymbol{p}, \tilde{\boldsymbol{p}}, u) \equiv \pi(\tilde{\boldsymbol{p}}) \left\{ 1 + \kappa \left[\alpha' \boldsymbol{p}(\lambda) + \tfrac{1}{2} \boldsymbol{p}(\lambda)' \boldsymbol{B} \boldsymbol{p}(\lambda) + \theta(\tilde{\boldsymbol{p}}, u) \right] \right\}^{1/\kappa},$$

---

where $e_i \equiv p_i q_i$ and $\boldsymbol{e} \equiv [e_1 \cdots e_n]'$ is the n-vector of expenditures on the goods $\boldsymbol{q}$. Similarly, it is shown in LaFrance (1985) that integrable, non-homothetic Linear Incomplete Demand Systems (LIDS) are characterized by linear identities among *quantities*, say,

$$(5) \qquad \boldsymbol{q} \equiv \alpha + \gamma (q_1 - \alpha_1)/\gamma_1 \ \forall \ (\boldsymbol{p}, m).$$

In this sense, case (a) closes the set of demand models that can be characterized by quantities demanded (LIDS), expenditures (LES), or budget shares (LA-AIDS) lying on a ray in n-dimensional space, regardless of the observed levels of prices and income.

where $\theta(\tilde{\boldsymbol{p}},u)$ is 0° homogeneous in $\tilde{\boldsymbol{p}}$ and increasing in u, but otherwise cannot be identified (LaFrance (1985); LaFrance and Hanemann (1989)). It also can be shown that the demands in (6) are homothetic, with income elasticities equal to $1-\kappa \; \forall \; \kappa \in \mathbb{R}$.

This simple procedure for nesting the homothetic LA-AIDS and LIDS within a homothetic PIGL-IDS easily generalizes to the non-homothetic, integrable AIDS,

$$(8) \qquad \boldsymbol{w} = \alpha + \boldsymbol{B}\ln(\boldsymbol{p}) + \gamma\left[\ln(m) - \alpha_0 - \alpha'\ln(\boldsymbol{p}) - \tfrac{1}{2}\ln(\boldsymbol{p})'\boldsymbol{B}\ln(\boldsymbol{p})\right].$$

To show this, we require a second result, which states that (8) is a special case of a complete class of incomplete demand models that can be characterized as follows. Let $y \equiv g_0(m)$ and $x_i \equiv g(p_i)$, $i = 1\ldots n$, where $g_0(\cdot)$ and $g(\cdot)$ are strictly increasing and twice continuously differentiable functions on $\mathbb{R}_{++}$, and write the n-vector inverse of $\boldsymbol{g}(\cdot)$ as $\boldsymbol{p}(\boldsymbol{x})$. Suppose that, after appropriate transformations, the demand functions for the n goods $\boldsymbol{q}$ can be written in terms of a linear function of $y(\boldsymbol{x},\tilde{\boldsymbol{p}},u) \equiv g_0\left(e(\boldsymbol{p}(\boldsymbol{x}),\tilde{\boldsymbol{p}},u)\right)$ and linear and quadratic functions of $\boldsymbol{x}$, with no interaction terms between $\boldsymbol{x}$ and y,

$$(9) \qquad \frac{\partial y(\boldsymbol{x},\tilde{\boldsymbol{p}},u)}{\partial x_i} = \alpha_i + \beta_i'\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{x}'\Delta_i\boldsymbol{x} + \gamma_i y(\boldsymbol{x},\tilde{\boldsymbol{p}},u),\; i = 1,\ldots,n$$

where, without loss in generality, $\gamma_1 \neq 0$ and each $n\times n$ matrix, Δ_i, is symmetric $\forall\; i$. Then we have the following.

Lemma 2. *The system of partial differential equations in* (9) *is integrable if, and only if, it can be written in the form*

$$\frac{\partial y(\boldsymbol{x},\tilde{\boldsymbol{p}},u)}{\partial \boldsymbol{x}} = \tilde{\alpha} + \tilde{\boldsymbol{B}}\boldsymbol{p} + \gamma\left[y(\boldsymbol{x},\tilde{\boldsymbol{p}},u) - \alpha_0 - \tilde{\alpha}'\boldsymbol{x} - \tfrac{1}{2}\boldsymbol{x}'\tilde{\boldsymbol{B}}\boldsymbol{x}\right]$$

where α_0 *is a scalar (that may be a function of other prices),* $\tilde{\alpha} = \alpha - \alpha_0\gamma$ *is an* $n\times 1$ *vector,* $\tilde{\boldsymbol{B}}$ *is a symmetric* $n\times n$ *matrix that satisfies* $\tilde{\boldsymbol{B}} = \boldsymbol{B} + \gamma\alpha'$,

$$where\ \boldsymbol{B} = [\beta_1 \cdots \beta_n],\ and\ \Delta_i = -\gamma_i \tilde{\boldsymbol{B}}\ \forall\ i.$$

Note that if $g_0(\cdot)$ and $g(\cdot)$ are both natural logarithmic functions we obtain the integrable AIDS. In addition, with the above definitions for $m(\kappa)$ and $\boldsymbol{p}(\lambda)$, we can write an integrable non-homothetic PIGL-IDS that is linear in the Box-Cox expenditure term and linear and quadratic in the Box-Cox price vector as,

$$\text{(10)} \qquad \boldsymbol{w} = m^{-\kappa}\boldsymbol{P}^{\lambda}\left\{\alpha + \boldsymbol{Bp}(\lambda) + \gamma\left[m(\kappa) - \alpha_0 - \alpha'\boldsymbol{p}(\lambda) - \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{Bp}(\lambda)\right]\right\},$$

where, for notational simplicity, we have dropped the tildes over the parameters.

Unlike the homothetic case, for all (κ, λ) pairs, this flexible functional form allows one to estimate the income aggregation function through the Box-Cox parameter κ. If $\kappa = 0$ we obtain the integrable AIDS model, if $\kappa = 1$ we obtain the linear-quadratic IDS (LQ-IDS) of LaFrance (1990), and for all (κ, λ) pairs we obtain an integrable PIGL-IDS.[6] Finally, it can be shown that the expenditure function for (10) is

$$\text{(11)} \qquad \mathrm{e}(\boldsymbol{p}, \tilde{\boldsymbol{p}}, u) \equiv \pi(\tilde{\boldsymbol{p}})\left\{1 + \kappa\left[\alpha_0 + \alpha'\boldsymbol{p}(\lambda) + \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{Bp}(\lambda) + \theta(\tilde{\boldsymbol{p}}, u)e^{\gamma'\boldsymbol{p}(\lambda)}\right]\right\}^{1/\kappa},$$

where $\theta(\cdot)$ has the same properties as for the expenditure function of equation (7). Note that the expenditure function (11) simply generalizes the one in (7) via the additional term α_0, which is often fixed at zero in empirical applications, and the factor $e^{\gamma'\boldsymbol{p}(\lambda)}$, which produces the n-vector of non-homothetic coefficients in the demand model.

This nesting procedure also generalizes to demand models that include linear and quadratic terms in the Box-Cox transformation of deflated income (QES-IDS). In order to motivate and demonstrate this, suppose that we have a transformed demand system of the form

---

[6] See Agnew (1998) for a comprehensive development and application of this full rank two PIGL-IDS.

(12) $$\frac{\partial y(\boldsymbol{x},u)}{\partial \boldsymbol{x}} = \sum_{i=0}^{k} \boldsymbol{\alpha}_i(\boldsymbol{x}) g_i\left(y(\boldsymbol{x},u)\right).$$

In a remarkable paper, by setting $y = \ln(m)$ and $\boldsymbol{x} = \ln(\boldsymbol{p})$, Gorman (1981) showed three things about all complete demand systems of this type:

(i) After normalizing for a unique representation and to account for adding up, making a change of variables, and accounting for some of the implications of Slutsky symmetry, the nonlinear partial differential equations can be transformed into a set of homogeneous linear ordinary differential equations in functions of the natural logarithm of income. From the theory of differential equations, solutions to this system are of the form $h_i(m) = m^{\lambda_i}\left(\ln(m)\right)^i$, where each λ_i is a root of the characteristic polynomial for the linear ordinary differential equations. In general, such characteristic roots can be either real or complex, and complex roots come in conjugate pairs that may have both real and complex parts.

(ii) However, Gorman was able to show that if the rank of the $n\times k$ coefficient matrix $\boldsymbol{A}(\boldsymbol{x}) \equiv [\alpha_{ij}(\boldsymbol{x})]$ equals at least three, then: (a) the characteristic roots are either purely real or purely complex (i.e., all roots of the form $\lambda_i = a_i + b_i\sqrt{-1}$ must have $a_i = 0$ if $b_i \neq 0$ and conversely, $b_i = 0$ if $a_i \neq 0$); (b) if any roots are real, there are no complex roots, and conversely; and (c) for real roots, there are no product terms of the form $m^{\alpha}\left(\ln(m)\right)^{\beta}$ with both $\alpha \neq 1$ and $\beta \neq 0$.

(iii) Finally, Gorman showed that the rank of $\boldsymbol{A}(\boldsymbol{x})$ is at most equal to three.

For rank three demand systems, this completely specifies the class of functional forms for the expenditure terms. Only three mutually exclusive cases are possible: (a) $m\left(\ln(m)\right)^r$, where each r is an integer; (b) $m^{1+\kappa}$, where κ may or may not be an integer; and (c) $m\sin\left(r\ln(m)\right)$ and $m\cos\left(r\ln(m)\right)$, for some $r \geq 0$, with both sine and cosine terms appearing as a conjugate complex pair. In other words, for rank three demand systems, the model must take one of the following three forms:

(13) $$q = \alpha_0(x)m + \sum_{j=1}^{k} \alpha_j(x)m\left(\ln(m)\right)^j \; ;$$

(14) $$q = \alpha_0(x)m + \sum_{\kappa \in T} \beta_\tau(x)m^{1-\kappa} + \sum_{\kappa \in T} \gamma_\tau(x)m^{1+\kappa} \; ,$$

where T is a set of nonzero constants; or

(15) $$q = \alpha_0(x)m + \sum_{\tau \in T} \beta_\tau(x)m\sin\left(\tau\ln(m)\right) + \sum_{\tau \in T} \gamma_\tau(x)m\cos\left(\tau\ln(m)\right),$$

where T is a set of positive constants. The case defined by (13) includes Muellbauer's (1975) PIGLOG model and extensions that are polynomials in $\ln(m)$, while the case given by (14) includes polynomials in income, as well as Muellbauer's PIGL model and extensions that are polynomials in m^κ.

Demand models that have *full rank* are characterized by the property that the rank of the matrix $A(x)$ is equal to the number of its columns, that is, the number of different income functions, $g_j(y)$. Clearly, full rank one demand models must be homothetic,

(16) $$q = \alpha_0(x)m,$$

due to adding up. Muellbauer (1975) showed that all full rank two demand systems are either PIGL or PIGLOG, that is, either

(17) $$q = \alpha_0(x)m + \alpha_1(x)m^{1+\kappa},$$

for some $\kappa \neq 0$, or

(18) $$q = \alpha_0(x)m + \alpha_1(x)m\ln(m).$$

Muellbauer's results are based on two things. First, adding up requires that one of the $g_i(\cdot)$ is identically m. Second, symmetry requires the system of partial differential equations to be linear in some transformation of income. Only two transformations satisfy both of these conditions. One is the Bernoulli equation,

(19) $$\frac{\partial e(\boldsymbol{p},u)^{\kappa}}{\partial \boldsymbol{p}} = \kappa \cdot e(\boldsymbol{p},u)^{\kappa-1} \cdot \left(\frac{\partial e(\boldsymbol{p},u)}{\partial \boldsymbol{p}} \right) = \beta_0(\boldsymbol{p}) + \beta_1(\boldsymbol{p}) e(\boldsymbol{p},u)^{\kappa}.$$

which has the PIGL form

(20) $$\frac{\partial e(\boldsymbol{p},u)}{\partial \boldsymbol{p}} = \beta_0(\boldsymbol{p}) e(\boldsymbol{p},u)^{1-\kappa} + \beta_1(\boldsymbol{p}) e(\boldsymbol{p},u).$$

The second is the logarithmic transformation,

(21) $$\frac{\partial \ln\left(e(\boldsymbol{p},u)\right)}{\partial \boldsymbol{p}} = \left(\frac{\partial e(\boldsymbol{p},u)/\partial \boldsymbol{p}}{e(\boldsymbol{p},u)} \right) = \beta_0(\boldsymbol{p}) + \beta_1(\boldsymbol{p}) \ln\left(e(\boldsymbol{p},u)\right),$$

which has the PIGLOG form

(22) $$\frac{\partial e(\boldsymbol{p},u)}{\partial \boldsymbol{p}} = \beta_0(\boldsymbol{p}) \cdot e(\boldsymbol{p},u) + \beta_1(\boldsymbol{p}) \cdot e(\boldsymbol{p},u) \cdot \ln\left(e(\boldsymbol{p},u)\right).$$

Gorman (1981) conjectured that polynomials of order two are the most general nondegenerate cases for rank three demand systems. Following up on this conjecture by exploiting the methods of Van Daal and Merkies (1989), Lewbell (1990) showed that all full rank three generalizations of the PIGL and PIGLOG models are exact analogues to the QES. Lewbell (1990) also derived a rank three trigonometric model of the form (15), although we do not make use of that result here.

All of the above results on the rank of the (price dependent) coefficient matrix and the functional form of the income terms in aggregable demand models rely on the adding up condition for a complete system of demand equations. In this study, however, we are interested in the aggregate U.S. demand for food items, and apply an incomplete demand system approach along the lines of Gorman (1965), Epstein (1982), LaFrance (1985) and

LaFrance and Hanemann (1989). Since total food expenditures make up only a small part of a typical household's budget, the budget identity takes the form of a strict inequality. In addition, while we have a rich and long time-series data set on consumption and prices for individual food items, we do not have such detailed information on the individual consumption levels or prices of other goods. These considerations preclude us from directly applying the above results in this study.

Nevertheless, even for incomplete demand systems of the form (12), it can be shown that Gorman's rank theorem is a corollary to symmetry for polynomials in income, PIGL and PIGLOG functional forms.

Lemma 3. *If* $y \equiv g_0(m)$ *is m,* m^{κ}*, or ln(m) and the (possibly incomplete) system of demands (12) is integrable, then there exist real-valued functions,* $\varphi_i : \mathbb{R}^n_+ \to \mathbb{R},\ i = 2,\ldots,k$ *such that*

$$\alpha_i(\boldsymbol{x}) \equiv \varphi_i(\boldsymbol{x})\alpha_k(\boldsymbol{x}) \ \forall\ i \geq 2\,.$$

Therefore, we can proceed with our nesting procedure by extending the rank two PIGL-IDS expenditure function in (11) to one that is, at least possibly, rank three and that generates a relatively simple form for the quadratic terms in the demand equations. A simple, and convenient, choice is a *quasi-indirect utility function* (Hausman (1981); LaFrance (1985); LaFrance and Hanemann (1989)) that can be written in a form that is (in principle) consistent with the QES originally developed in Howe, Pollak and Wales (1979),[7]

$$\text{(23)} \qquad \varphi(\boldsymbol{p}, m) = -\left\{ \frac{1}{\left[m(\kappa) - \alpha_0 - \alpha' \boldsymbol{p}(\lambda) - \frac{1}{2}\, \boldsymbol{p}(\lambda)' \boldsymbol{B} \boldsymbol{p}(\lambda) \right]} + \delta' \boldsymbol{p}(\lambda) \right\} e^{\gamma' \boldsymbol{p}(\lambda)}\,.$$

Applying the methodology of LaFrance and Hanemann (1989), it can be shown that (23) is equivalent to an expenditure function of the form

$$(24)\quad e(\boldsymbol{p},\tilde{\boldsymbol{p}},u)\equiv \pi(\tilde{\boldsymbol{p}})\left\{1+\kappa\left[\alpha_0+\boldsymbol{\alpha}'\boldsymbol{p}(\lambda)+\tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)-\frac{e^{\gamma'\boldsymbol{p}(\lambda)}}{\left(\delta'\boldsymbol{p}(\lambda)e^{\gamma'\boldsymbol{p}(\lambda)}+\theta(\tilde{\boldsymbol{p}},u)\right)}\right]\right\}^{1/\kappa}.$$

That is, the QPIGL-IDS expenditure function in (24) generalizes the non-homothetic PIGL-IDS expenditure function in (11) by replacing the term $\theta(\tilde{\boldsymbol{p}},u)e^{\gamma'\boldsymbol{p}(\lambda)}$ with the term $-[\delta'\boldsymbol{p}(\lambda)+\theta(\tilde{\boldsymbol{p}},u)e^{-\gamma'\boldsymbol{p}(\lambda)}]^{-1}$, which produces the *n*-vector of parameters δ associated with the quadratic term in supernumerary income, in addition to the *n*-vector of parameters γ associated with the linear supernumerary income terms.

Finally, an application of Roy's identity to (23) generates a QPIGL-IDS in budget share form as

$$(25)\qquad \boldsymbol{w}=m^{-\kappa}\boldsymbol{P}^{\lambda}\left\{\alpha+\boldsymbol{B}\boldsymbol{p}(\lambda)+\gamma\left[m(\kappa)-\alpha_0-\alpha'\boldsymbol{p}(\lambda)-\tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)\right]\right.$$
$$\left.+\left[\boldsymbol{I}+\gamma'\boldsymbol{p}(\lambda)\right]\delta\left[m(\kappa)-\alpha_0-\alpha'\boldsymbol{p}(\lambda)-\tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)\right]^2\right\}.$$

Assuming that α and **B** do not completely vanish simultaneously, it follows that: (a) γ ≠ **0,** δ ≠ **0** is necessary and sufficient for a full rank three QPIGL-IDS; (b) γ ≠ **0,** δ = **0** is necessary and sufficient for a full rank two, non-homothetic PIGL-IDS; (c) γ = **0,** δ ≠ **0** is necessary and sufficient for a full rank two QPIGL-IDS that excludes the linear term; and (d) γ = δ = **0** is necessary and sufficient for a homothetic PIGL-IDS. Thus, we obtain a

---

[7] Solve (11) for θ, transform to $\tilde{\theta}=-1/\theta$ to get $\tilde{\varphi}(\boldsymbol{p},m)=-\,e^{\gamma'\boldsymbol{p}(\lambda)}\big/[m(\kappa)-\alpha_0-\alpha'\boldsymbol{p}(\lambda)-\tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)]$, and add the term $-\delta'\boldsymbol{p}(\lambda)e^{\gamma'\boldsymbol{p}(\lambda)}$ to obtain (23).

rich class of models that permits nesting, testing and estimating the rank and functional form of the income aggregation terms in incomplete demand systems.

3. Estimating the U.S. Income Distribution

When a demand model is nonlinear in income, the demand equations do not aggregate directly across individual decision units to average (per capita or per family) income at the market level. The advantage of the Gorman class of Engel curves is that, when information on the income distribution across economic units is available, only a small number of summary statistics from this distribution are required to obtain a theoretically consistent, aggregable demand model. Indeed, all full rank three Gorman Engel curve demand models require three summary statistics from the income distribution, e.g., a QPIGL requires the cross-sectional means of $m_h^{1-\kappa}$, m_h, and $m_h^{1+\kappa}$. To calculate these means, however, we need information on the distribution of income.

The U.S. Bureau of the Census publishes annually quintile ranges, intra-quintile means, the top five-percentile lower bound for income, and the mean income within the top five-percentile range for all U.S. families. These data are currently available for 1947-1998 on the U.S. Bureau of the Census World Wide Web site, and for the years 1929, 1935/36, 1941, 1944 and 1946 from the Census Bureau's historical statistics (U.S. Department of Commerce, 1972). Several issues arise regarding the use of these data to estimate the U.S. income distribution. First and perhaps foremost is an appropriate methodology for obtaining a reasonable density function given the probability ranges and intra-range means. In this paper, we consider three possibilities, depicted in figure 1 for 1997, which are developed and explained in this section.

The simplest, most naïve and uninformative approach is to construct a sequence of piecewise uniform densities on each of the first four quintile ranges, the 85-95 percentile range, and the top five percentile range.[8] However, these piecewise uniform densities

---

[8] The mean for the 80-95 percentile range is calculated as $\mu_{80-95\%} = (.20\mu_{80-100\%} - .05\mu_{95-100\%})/.15$. The 85-95 percentile range is the interval from the lower limit of the top quintile to the lower limit of the top five percentile range, while the top five percentile mean is assumed to be the midpoint of that range for the piecewise uniform densities discussed in this subsection.

generally do not satisfy the intra-quintile and top five percentile mean conditions. A slightly more informative solution is to construct a pair of uniform densities on each range, separated at the intra-range mean, and with total probabilities that sum to .20, .15, or .05, as appropriate. To illustrate, let $[m_{i-1}, m_i)$ denote the i^{th} income range, μ_i the i^{th} intra-range mean, and p_i the proportion of the total number of U.S. families whose incomes that fall within this range. We calculate a piecewise uniform density on $[m_{i-1}, m_i)$, satisfying

(26) $$f(x) = \begin{cases} f_{1,i}, \ \forall \ x \in [m_{i-1}, \mu_i) \\ \\ f_{2,i}, \ \forall \ x \in [\mu_i, m_i) \end{cases},$$

subject to the probability and mean conditions,

(27) $$f_{1,i}(\mu_i - m_{i-1}) + f_{2,i}(m_i - \mu_i) = p_i \ ,$$

(28) $$\frac{f_{1,i}(\mu_i^2 - m_{i-1}^2) + f_{2,i}(m_i^2 - \mu_i^2)}{2 p_i} = \mu_i \, .$$

Solving the two constraint equations for the two density levels gives

(29) $$f_{1,i} = \left(\frac{m_i - \mu_i}{\mu_i - m_{i-1}} \right) \frac{p_i}{(m_i - m_{i-1})},$$

(30) $$f_{2,i} = \left(\frac{\mu_i - m_{i-1}}{m_i - \mu_i} \right) \frac{p_i}{(m_i - m_{i-1})}.$$

This density is illustrated in figure 1 for 1997 by the series of horizontal line segments.

Of course, this density estimator is *ad hoc*, discontinuous at eleven points,[9] an artifact of the manner in which the data are reported, and does not satisfy any particular validating criterion. Moreover, since we have a fixed number of observations in each year on quintile limits, intra-quintile means, and the top five percentile lower limit and mean, we cannot appeal to properties like consistency as "sample size" increases. Therefore, a density estimator motivated by some formalism is preferable to this piecewise uniform density. Two possible approaches to this issue are considered here.

The first is the method of maximum entropy, which is based on information theory and well-known to possess several desirable properties (Zellner 1988). This approach generates an income density that is smooth and monotone within the pre-specified income ranges and satisfies each probability and intra-range mean condition exactly, but is discontinuous at the boundary between each pair of contiguous income ranges. For 1997, this density is depicted in figure 1 by the series of piecewise exponential curves marked with solid black circles.

The second approach is a parametric, truncated three-parameter lognormal density. This density is smooth everywhere and has a general shape that is similar to the MAXENT density, but does not satisfy either the probability or mean conditions exactly in any range of income. For 1997, this density is depicted in figure 1 by the smooth curve marked with empty circles.

3.1 The Maximum Entropy Density

The starting point for this estimator for the income distribution is the maximum entropy (MAXENT) criterion,

---

[9] For simplicity, and lack of a better alternative, the intra-range mean of the top five percentile group is assumed to be located at the center of that range, making the top percentile uniform density continuous up to the point $x_{.95} + 2\mu_{.95}$, which reduces the number of discontinuities from twelve to eleven.

(31) $$E \equiv \max_{\{f\}} -\int_0^\infty f(x)\ln\left(f(x)\right)dx,$$

to select a density function, $f(x)$, that satisfies, in this case, the piecewise probability and mean constraints for each quintile and the top five percent of family incomes,

(32) $$\int_{m_{i-1}}^{m_i} f(x)dx = p_i,\ i = 1,...,6,$$

(33) $$\int_{m_{i-1}}^{m_i} xf(x)dx = p_i\mu_i,\ i = 1,...,6,$$

where $m_0 = 0$, $m_6 = \infty$, and $p_i = .20,\ i = 1,...,4,$ $p_5 = .15$, and $p_6 = .05$.

The MAXENT criterion is motivated, developed and discussed in detail in Csiszár (1991), Gokhale and Kullback (1978), Jaynes (1957a, 1957b, 1984), Kullback (1959) and Shannon (1948), while detailed derivations and discussions of the MAXENT are contained in Tobias and Zellner (1997), Zellner (1988, 1997) and Zellner, Tobias, and Ryu (1997). In the present case, the MAXENT density is the (unique) proper density on ⊕+ that minimizes the average logarithmic height of the density function while satisfying the integration and moment constraints, which in turn provide the available prior information in this problem. In this respect, it is the density function that is "closest" to the piecewise uniform density function originally considered in this section, where close in this context is defined by the Kullback-Leibler cross entropy pseudo-distance measure (Golan, Judge, and Miller (1996)).

Formally, the MAXENT density is the maximum with respect to {f} and minimum with respect to the Lagrange multipliers $\{\vec{\lambda}_0, \vec{\lambda}_1)$ in the Lagrangean function,

(34) $$\ell = -\int_0^\infty f(x)\ln[f(x)]dx + \sum_{i=1}^{6}\lambda_{0,i}\left[p_i - \int_{m_{i-1}}^{m_i} f(x)dx\right] + \sum_{i=1}^{6}\lambda_{1,i}\left[p_i\mu_i - \int_{m_{i-1}}^{m_i} xf(x)dx\right]$$

$$= -\sum_{i=1}^{6}\int_{m_{i-1}}^{m_i} f(x)\left\{\ln[f(x)] + \lambda_{0,i} + \lambda_{1,i}x\right\}dx + \sum_{i=1}^{6} p_i\left(\lambda_{0,i} + \lambda_{1,i}\mu_i\right).$$

Note that the second line of (34) involves a sum of integrals in {*f*} with no differential or boundary value constraints on the optimal density function. It therefore follows from optimal control theory (see, e.g., Seierstad and Sydsæter, 1987, or Clegg, 1968) that, for any choice of the Lagrange multipliers on the isoperimetric (i.e., integral) constraints, the necessary condition for a maximum of ℓ with respect to {*f*} is to find the pointwise maximum of each of the six individual integrands with respect to *f*. Also, note that for each term in this sum, the integrand is strongly concave in *f*, while the constraints are linear in {*f*} so that this maximum is unique.

Thus, the first-order conditions for a constrained maximum with respect to {*f*} are

(35) $$-\left\{1 + \ln[f(x)] + \lambda_{0,i} + \lambda_{1,i}x\right\} = 0,\ \forall\ x \in [m_{i-1}, m_i),\ i = 1,\ldots,6,$$

together with the integral constraint conditions (32) and (33). Solving (35) for *f* for each $x \in [m_{i-1}, m_i),\ i = 1,\ldots,6$, then implies

(36) $$f(x) = \exp\left\{-\left(1 + \lambda_{0,i} + \lambda_{1,i}x\right)\right\},\ \forall\ x \in [m_{i-1}, m_i),\ i = 1,\ldots,6,$$

so that we obtain a sequence of six exponential densities, each defined on their respective percentile ranges. Next, integrating (36) over each percentile range, substituting the resulting the right-hand side expression into the associated probability constraint (32), and solving for $e^{-(1+\lambda_{0,i})}$ gives

(37) $$f(x) = \begin{cases} -p_i\lambda_{1,i}e^{-\lambda_{1,i}x} \Big/ \left(e^{-\lambda_{1,i}m_i} - e^{-\lambda_{1,i}m_{i-1}}\right), & x \in [m_{i-1}, m_i),\ i = 1,\ldots,5, \\ p_6\lambda_{1,6}e^{-\lambda_{1,6}(x - m_5)}, & x \in [m_5, \infty). \end{cases}$$

with p_i = 0.20, i = 1…4, p_5 = 0.15, and p_6 = 0.05. Finally, multiplying x times the right-hand side of (37), integrating again over each percentile range, substituting the resulting right-hand side expression into the associated mean constraint (33), and rearranging terms gives the defining equations for the Lagrange multipliers for the mean constraints in the form

(38) $$e^{\lambda_{1,i}(m_i - m_{i-1})} - \left[\frac{1+\lambda_{1,i}(m_i - \mu_i)}{1-\lambda_{1,i}(\mu_i - m_{i-1})}\right] = 0,\ i = 1,\ldots,5,$$

(39) $$\lambda_{1,6} = 1/(\mu_6 - m_5).$$

It is self evident from (39) that $\lambda_{1,6}$ is strictly positive and can be readily calculated from the data, while there is no closed form expression for the solution to (38). However, it is straightforward to calculate the unique solutions for each of the first five $\lambda_{1,i}$ terms from the following set of elementary observations. First, if $\mu_i = (m_i + m_{i-1})/2$, then there is only one solution to (38), namely, $\lambda_{1,i}$ = 0. Second, the first term in (38) is always positive, while the second term has a zero at $\lambda_{1,i} = -1/(m_i - \mu_i) < 0$ and a pole at $\lambda_{1,i} = 1/(\mu_i - m_{i-1}) > 0$. It follows that if $\mu_i > (m_i + m_{i-1})/2$, then $\lambda_{1,i}$ < 0 and a simple interval halving procedure over $\lambda_{1,i} \in (-1/(m_i - \mu_i), 0)$ produces an approximate solution with accuracy on the order of 2^{-n}, where n is the number of iterations, or interval squeezes. Similarly, if $\mu_i < (m_i + m_{i-1})/2$, then $\lambda_{1,i}$ > 0 and a successive interval halving procedure on $\lambda_{1,i} \in (0, 1/(\mu_i - m_{i-1}))$ produces the desired approximate solution. In every case in the empirical application for this paper, the numerical estimates of $\lambda_{1,i}$ stopped changing (in double precision, to fifteen digits) after less than twenty iterations.

Figure 1 illustrates the MAXENT solution for the U.S. Income Distribution in 1997. As seen in the figure, this solution is made up of piecewise exponential densities marked with solid circles, with a discontinuity at the boundary between each pair of contiguous percentile ranges. Similar to the mean constrained piecewise uniform density in the figure, the MAXENT density intersects the vertical axis above zero, is unimodal and is skewed the right. These properties of the piecewise uniform and MAXENT densities – piecewise continuous, a single mode, and positively skewed – are maintained in all years of the available time series data. The number of discontinuities in the MAXENT density (five) is less than half of the number in the piecewise uniform density (eleven). With the frequent exception of the boundary between the first and second quintile, the absolute magnitude of the discontinuities in the MAXENT density at the percentile boundaries tends to be substantially less than in the piecewise uniform. Intuitively at least, these two properties, in addition to the formal motivation and justification on information theoretic grounds, is a significant advantage for the MAXENT density relative to a mean constrained piecewise uniform density. Hence, we do not employ the piecewise uniform density in the empirical demand analysis discussed in the next section.

However, the number and location of the discontinuities in the MAXENT solution remain as artifacts of the way in which the income distribution data is reported by the Census Bureau.[10] The ultimate goal of this study is to estimate the coefficients of a demand model that aggregates across income by using the income distribution that we construct from a limited set of available prior information. The discontinuities in the piecewise exponential MAXENT solution beg for a continuous alternative that can be used to estimate and evaluate the sensitivity of the demand model's parameter estimates and the conclusions drawn from those estimates. The specific parametric alternative

[10] It is possible to construct a continuous MAXENT solution for the income distribution wherein the first-order derivative of the MAXENT density (with respect to x) is an absolutely bounded control variable in the optimization problem. One possible (although arbitrary) choice for the maximal absolute slope of this piecewise differentiable MAXENT density is $\sup\left|f_0'(x)\right|$, where $f_0(x)$ is the discontinuous MAXENT solution. It can be shown that this piecewise differentiable solution is characterized by a sequence of piecewise exponential and linear segments, with the boundary point between each pair of contiguous segments determined endogenously as part of the MAXENT optimization problem. One also could go a step further and constrain the second-order derivative of the density to be the (absolutely bounded) control variable, which generates a ∂^1 sequence of exponential and quadratic splines. However, constructing either of these modified MAXENT solutions involves finding a suitable set of roots for a fairly large set of highly nonlinear equations, each of which has multiple zeroes. As a result, we do not pursue a continuity or differentiability constrained MAXENT solution in this paper.

considered in this paper is a truncated three-parameter lognormal distribution, which is motivated, derived and discussed next.

3.2 The Truncated Three-Parameter Lognormal Distribution

Let the random variable z have a standard normal distribution and define the latent income variable x by $\ln(x-\theta)=\mu+\sigma z$, where $x>\theta$ and $\{\mu,\sigma,\theta\}$ are constants. Then the random variable x has a *three-parameter lognormal distribution*, with probability density function (*pdf*) defined by

(40) $$f(x;\mu,\sigma,\theta)=\frac{1}{\sqrt{2\pi}\sigma(x-\theta)}\exp\left\{-\frac{1}{2\sigma^2}\left[\ln(x-\theta)-\mu\right]^2\right\}.$$

A few motivations for this specification in the current study include the following:

(a) As the above discussion clarifies, the piecewise uniform distribution and MAXENT distributions, which are both constructed to satisfy the quintile and top five-percentile probability and conditional mean conditions, are unimodal, skewed to the right, and intersect the vertical axis in the income/pdf plane at strictly positive values.

(b) Both the piecewise uniform and MAXENT density functions are discontinuous at arbitrary points defined by the construction and reporting procedures of the U.S. Bureau of the Census. We expect *a priori* that the income distribution would be better approximated by a continuous density function.

(d) The two-parameter lognormal distribution has been criticized for estimating income distributions (McDonald, 1984), and in the present case is incompatible with a positive intercept for the *pdf*. And,

(e) Combined with a simple, accurate, closed-form and invertible approximation to the standard normal cumulative distribution function (*cdf*), which is discussed in detail below, the truncated three-parameter log-normal distribution is simple to estimate and gives excellent results, given the limited nature of the available prior information on the U.S. income distribution over time.

Since we only observe values of $x \geq 0$, we are interested in estimating the conditional distribution of x given that $x \geq 0$. Therefore, let $x_0 \equiv 0$, define the standardized zero income limit by $z_0 = (\ln(-\theta) - \mu)/\sigma$, and denote the standard normal *cdf* at z_0 by $\Phi(z_0) = \int_{-\infty}^{z_0} \varphi(z)dz$, where $\varphi(z) = (1/\sqrt{2\pi})e^{-z^2/2}$ is the standard normal *pdf*. Note that z_0, and hence $\Phi(z_0)$, depends on the parameters $\{\mu,\sigma,\theta\}$. Then the conditional *pdf* for x given $x \geq 0$ is defined by

$$(41) \qquad f(x|x \geq 0; \mu, \sigma, \theta) = \frac{1}{\sqrt{2\pi}\sigma(x-\theta)(1-\Phi(z_0))} \exp\left\{-\frac{1}{2\sigma^2}\left[\ln(x-\theta)-\mu\right]^2\right\}.$$

This is the *pdf* of interest. However, the information we have for each year takes the form of income limits, x_i, where $\Pr(x \leq m_i \mid x \geq 0) = F_i$ = .2, .4, .6, .8, .95 for i = 1, 2, 3, 4, 5, respectively. Therefore, to estimate the truncated three-parameter lognormal distribution, in addition to z_0 above, define the standardized limits by

$$(42) \qquad z_i \equiv \frac{\ln(m_i - \theta) - \mu}{\sigma},\ i = 1,\ldots,5,$$

and the empirical limit equations,

$$(43) \qquad F_i = \Phi(z_i)/(1-\Phi(z_0)) + \varepsilon_i,\ i = 1,\ldots,5,$$

where ε_i is an estimation error. This produces five observations in the three-parameters $\{\mu,\sigma,\theta\}$, which can be estimated by nonlinear least squares.[11]

---

[11] We also could generate an additional six data points for each year based on the intra-percentile mean conditions. In practice, however, adding these additional data points created numerical overflows due to positive-valued estimates for the parameter θ in years prior to 1971. As a result, only the five data points associated with the probability constraints are considered here. However, the points estimates for {θ,μ,σ} are virtually identical using all eleven observations per year in each year over the period 1971-1999 as when only the percentile range probability conditions are used. See Johnson, Kotz and Balakrishnan (1994), chapter 14, for additional discussions on the numerical difficulties typically encountered when estimating a three-parameter lognormal density function.

However, due to the small number of observations available in each year, rather than using numerical integration to evaluate the standard normal *cdf* multiple times at each observation in each iteration of the nonlinear squares estimation procedure, we use the following simple, accurate and invertible approximation to the standard normal *cdf*. For any compact set of values of the standard normal variate, z, a variant of Taylor's theorem[12] implies that a polynomial of sufficiently high order in the exponential term of a logistic-type *cdf* will closely approximate the standard normal *cdf*. In addition, symmetry of the normal distribution implies that all even powers in z for this polynomial expansion must be zero. In particular, $\forall\ z \in [-3,+3]$, we find

(44) $$\Phi(z) \approx \frac{e^{1.6092\cdot z+.066826\cdot z^3}}{\left(1+e^{1.6092\cdot z+.066826\cdot z^3}\right)},$$

with $R^2 = 1.00000$ and $\max_i \left|\hat{\varepsilon}_i\right| = 1.45\times10^{-5}$, using 0.001 increments for z. Plots of the *pdf* and *cdf* for the third-order logistic approximation are indistinguishable from those for the standard normal distribution.[13] The resolution is substantially better for the log-odds ratio of the standard normal and the polynomial $1.6092\cdot z + .066826\cdot z^3$. This is illustrated in

---

[12] Taylor's theorem is applied to the log-odds ratio $\ln\left[\Phi(z)/(1-\Phi(z))\right] = \beta_1 z + \beta_3 z^3 + o(z^4)$. Note that the log-odds ratio is an analytic function of z (i.e., has an infinite radius of convergence for the series $\forall\ z \in$ ⊕). By symmetry of the normal distribution, the error of approximation is $o(z^4)$. Terms can be easily added to the expansion in (6) to increase the range and precision of the approximation. However, numerical overflows are likely to be encountered (even using double precision) for $|z| > 5$. In the present case, the largest absolute value of z is associated with z_0, which is on the order of 2.0 – 2.5, and a 3rd order expansion is extremely accurate within this range.

[13] One significant advantage of this approximation to the n(0,1) *cdf*, though not relevant in the present study, is the fact that Cardano's formula can be used to invert Φ to obtain z in the form $z = u + v$, the only real root to the characteristic 3rd-order polynomial, with u and v defined (in closed form) by

$$u=\sqrt[3]{\ln\left(\frac{\Phi}{1-\Phi}\right)+\tfrac{1}{2}\sqrt{\frac{\Delta}{27}}}\ ,\quad v=\sqrt[3]{\ln\left(\frac{\Phi}{1-\Phi}\right)-\tfrac{1}{2}\sqrt{\frac{\Delta}{27}}}\ ,$$

where $\Delta = 5.5854\cdot10^5 + 6.0461\cdot10^3\cdot[\ln\{\Phi/(1-\Phi)\}]^2 > 0\ \forall\ \Phi\in(0,1)$. This should prove particularly useful for many types of simulations and Monte Carlo experiments, since only a single uniform (0,1) pseudo-random variate is required to obtain a pseudo-normal variate, and the support of the resulting pseudo-normal variate is compact (though empirically not very much different from a standard normal random variable).

figure 2, where it can be seen that the interval [-3, +3] can be used with virtually no error. Moreover, the probabilities in the remaining tails and hence the differences between those probabilities are very small and the approximation is excellent throughout ⊕.

3.3 Extrapolating the Income Data

Data for U.S. food consumption and retail prices, as well as additional variables that are described in the next section, have been obtained from LaFrance (1999a) for the years 1918–1995. However, observations on the Census Bureau's summary data for the income distribution are available for 1929, 1935/36, 1941 and 1946–98. One issue that arises in using this data in an aggregate U.S. food demand model, then, centers on predicting or extrapolating this income data for the years 1918–1928, 1930–40, 1942–43, and 1945. We take the following approach to address this question.

To forecast the lower limit of the second quintile (equivalently, the upper limit of the first quintile, $m_{1,t}$), we utilize data on per capita disposable personal income and the unemployment rate as predictors. We estimate a least squares relationship with the log of the second quintile lower limit as the dependent variable and a constant term, the log of average per capita disposable income and this variable squared, and the unemployment rate as regressors, with first-order serial correlation in the error terms,

$$\text{(45)} \qquad \ln(m_{1,t}) = \alpha + \beta_1 \ln(\mu_{pc,t}) + \beta_2 [\ln(\mu_{pc,t})]^2 + \beta_3 u_t + e_{1,t}, \quad e_{1,t} = \rho e_{1,t-1} + \varepsilon_{1,t},$$

where $\mu_{pc,t}$ is per capita disposable personal income and u_t is the annual average U.S. unemployment rate. The sample period for this equation is 1946–97 due to the presence of autocorrelation. Predicted values for $\ln(m_{1,t})$ are calculated from this regression equation for the years 1918–1928, 1930–40, 1942–43, and 1945 with the observations on average per capita disposable income and the unemployment rate for each year in which data of the income distribution is not available.

For the larger income limits, we follow a recursive forecasting procedure in which an ordinary least squares prediction equation is estimated using a constant term and first- second- and third-order powers of the log of the closest smaller limit as regressors,

(45) $$\ln(m_{i,t}) = \alpha_i + \beta_{i1}\ln(m_{i-1,t}) + \beta_{i2}[\ln(m_{i-1,t})]^2 + \beta_{i3}[\ln(m_{i-1,t})]^3 + e_{i,t},$$

for i = 2, …, 5. The summary statistics in present table 1 and the plots of the observed data and regression curves in figure 3 suggest that these conditional prediction equations are quite precise. For each year in which income distribution data is not reported by the Census Bureau, a recursive forecast obtained from the above sequence of steps are to fill in the missing observations on the income distribution percentile limits.

We follow a similar procedure to forecast the intra-percentile means, beginning with the first quintile mean as a linear function of a constant term, the log of average per capita income and the square of this variable, and the unemployment rate, with first-order autocorrelation,

(46) $$\ln(\mu_{1,t}) = \alpha + \beta_1\ln(\mu_{pc,t}) + \beta_2[\ln(\mu_{pc,t})]^2 + \beta_3 u_t + v_{1,t}, \quad v_{1,t} = \rho v_{1,t-1} + \upsilon_{1,t}.$$

As for the first quintile upper limit regression, the sample period for the first quintile mean income regression is 1946–97 due to the presence of autocorrelation. Predicted values for $\ln(\mu_{1,t})$ are calculated for the years 1918–1928, 1930–40, 1942–43, and 1945 using data on average per capita disposable income and the unemployment rate. These predictions, in turn, are used in the following recursive prediction procedure for the remaining intra-percentile mean incomes.

For each intra-percentile mean income above the first quintile, we estimate an ordinary least squares prediction equation using a constant term and first- second- and third-order powers of the log of the nearest smaller intra-percentile mean, the log of average per capita disposable income, and the unemployment rate as regressors,

(47) $\ln(\mu_{i,t}) = \alpha_i + \beta_{i1}\ln(\mu_{i-1,t}) + \beta_{i2}[\ln(\mu_{i-1,t})]^2 + \beta_{i3}[\ln(\mu_{i-1,t})]^3 + \beta_{i4}\ln(\mu_{pc,t}) + \beta_{i5}u_t + v_{i,t}$,

for i = 2, …, 6. The summary statistics in table 2 and the plots of the observed and predicted values of each intra-percentile mean in figure 4 suggest that these conditional prediction equations also are very precise. As in the previous case, we use the predicted values obtained from the least squares regressions to fill in the missing observations on the U.S. income distribution.

4. Estimating the Nested QPIGL-IDS for U.S. Food Demand

The system of empirical nested QPIGL-IDS demand equations that we estimate for U.S. food consumption for the years 1918–1995, excluding 1942–1946, can be written in deflated expenditure form as

(48) $$\boldsymbol{e}_t = m_t^{1-\kappa}\boldsymbol{P}_t^{\lambda}\left\{\boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{B}\boldsymbol{p}_t(\lambda) + \gamma\left[m_t(\kappa) - \boldsymbol{p}(\lambda)'\boldsymbol{A}\boldsymbol{s}_t - \tfrac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda)\right]\right.$$

$$\left. + \left[\boldsymbol{I} + \gamma'\boldsymbol{p}_t(\lambda)\right]\delta\left[m_t(\kappa) - \boldsymbol{p}_t(\lambda)'\boldsymbol{A}\boldsymbol{s}_t - \tfrac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda)\right]^2\right\} + \varepsilon_t,\ t = 1, \ldots, T,$$

where $\boldsymbol{e}_t = [p_{1t}q_{1t} \ldots p_{nt}q_{nt}]'$ is the n-vector of deflated per family annual expenditures on individual food items, $\boldsymbol{s}_t$ is a vector that includes a constant, the mean, variance and skewness of the U.S. population's age distribution, the proportion of the U.S. population that is Black and the proportion of the population that is neither Black nor White, and ε_t is an n-vector of mean zero, identically distributed error terms. We specify the empirical model in expenditure form to keep all income terms on the right-hand side so that the mean values of all of the appropriate transformations of income are properly calculated across all U.S. families during the econometric estimation of the demand parameters.

Expanding the second line of equation (48) and grouping terms, the QPIGL-IDS demand equations can be rewritten in the form

$$(49)\qquad \boldsymbol{e}_t = \boldsymbol{P}_t^{\lambda}\left\{\left[\boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{B}\boldsymbol{p}_t(\lambda) - \gamma\left(\frac{1}{\kappa} + \boldsymbol{p}(\lambda)'\boldsymbol{A}\boldsymbol{s}_t + \frac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda)\right)\right.\right.$$

$$\left. + \left[\boldsymbol{I} + \gamma'\boldsymbol{p}_t(\lambda)\right]\delta\left(\boldsymbol{p}_t(\lambda)'\boldsymbol{A}\boldsymbol{s}_t + \frac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda) + \frac{1}{\kappa}\right)^2\right] m_t^{1-\kappa}$$

$$+ \frac{1}{\kappa}\left[\gamma - 2\left[\boldsymbol{I} + \gamma'\boldsymbol{p}_t(\lambda)\right]\delta\left(\boldsymbol{p}_t(\lambda)'\boldsymbol{A}\boldsymbol{s}_t + \frac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda) + \frac{1}{\kappa}\right)\right] m_t$$

$$\left. + \left[\frac{\boldsymbol{I} + \gamma'\boldsymbol{p}_t(\lambda)}{\kappa^2}\right]\delta\, m_t^{1+\kappa}\right\} + \varepsilon_t\,,$$

which shows explicitly that we require the means with respect to the income distribution of the variables $m_t^{1-\kappa}$, m_t, and $m_t^{1+\kappa}$ in each year to consistently estimate the demand parameters with aggregate market data. Estimation of the model's parameters therefore requires, for any given value of $\kappa \in (0, 1]$, numerical integration methods to evaluate the expected values of the three powers of income at each year in the sample period, where the expectation is taken over that year's estimated income distribution.

To accomplish this, we transform the positive half line $\oplus_+$ into the unit interval [0, 1) through a change of variables to $y = x/(1+|x|)$ and use Simpson's rule on a grid over the unit interval. To ensure that this numerical approximation is good, for the MAXENT distribution we compare the intra-quintile and the top five percentile mean incomes, which are known exactly, with the numerical values obtained through numerical integration. In this case, the numbers obtained by the two comparison methods were virtually identical. For the truncated three-parameter lognormal distribution, we use Monte Carlo simulation to estimate the income moments using a pseudo-random sample with 10^7 observations and compare the estimated values with the results obtained by numerical integration.[14]

We use two-step nonlinear seemingly unrelated regressions equations (NLSURE) estimation methods, combined with a one dimensional search over the income term's Box-Cox parameter κ. Only one iteration on the residual covariance matrix is calculated to avoid numerically over fitting one or more equations, which can occur with iterative NLSURE in large, highly parameterized demand models such as this.[15] A search over κ is used to incorporate the numerical integrations required to generate the aggregate income variables, which in turn depend upon the parameter κ. Symmetry of the coefficient matrix $\boldsymbol{B}$ is maintained throughout the estimation process in order to reduce the dimension of the parameter space from 527 to 317 estimated parameters.

Figure 5 presents the results of both steps of this estimation procedure, with the second step generated by searching over κ while holding the estimated residual covariance

---

[14] The Monte Carlo simulation proceeds as follows. First, a vector of $2N$ pseudo-random numbers is drawn from a uniform [0,1] distribution using a two-step compound linear congruential pseudo-random number generator. In the first step of this compound pseudo-random number generator, one number in the [0, 1] interval is drawn using one of seventeen mutually relatively prime linear congruential generators. The unit interval is broken up in into sixteen mutually exclusive and exhaustive sub-intervals of equal length, and the outcome from the first step determines which of the remaining sixteen other linear congruential generators is used to obtain a second pseudo-random number. That is, two pseudo-random numbers from the uniform [0, 1] distribution are drawn to obtain one useable observation. Mixing several relatively prime linear congruential pseudo-random number generators in this manner can be shown to overcome the known serial correlation problems inherent in a single linear congruential generator.

The second step is to use the Box–Mueller method to transform pairs of independent uniform (0, 1) observations into pairs of independent from the standard normal distribution. Each observation, z_i, from the sample of $2N$ *i.i.d.* N(0, 1) pseudo-random variables is transformed to an observation x_i from the three-parameter lognormal distribution using the formula $z_i = [\ln(x_i - \theta) - \mu]/\sigma$, equivalently, $x_i = \theta + \exp\{\mu + \sigma z_i\}$. If $x_i < 0$, then the observation is excluded from the truncated sample.

Finally, the sample moments of the retained observations on $x_i^{1-\kappa}$, x_i, and $x_i^{1+\kappa}$ are calculated using the number of retained observations as the divisor. Throughout the period 1918–95, this resulted in simulated sample sizes in the neighborhood of 9.8–9.9 × 10^6. Each simulated moment fell well within the range of 99–101% of the corresponding moment calculated by numerical integration.

[15] See, e.g., LaFrance (1999b), footnote 12 for a discussion of this issue. The crux of the matter is that all of the model parameters, which in the present case total 317, enter each of the demand equations, while there are only 76 time series observations. This creates a numerical possibility for a singular estimated covariance matrix when iterative NLSURE is employed, which generates an unbounded likelihood function.

matrix fixed at the value calculated from the total sum of squared residuals minimizing value for κ. As can be seen from the figure, this value is 1.00 for the truncated three-parameter lognormal (T3PLN) density and 1.03 for the MAXENT density. Conversely, the optimal values for κ obtained in the second iteration of the NLSURE procedure are 1.03 and 1.00 for the T3PLN and MAXENT income densities, respectively.

There are several important aspects of the results that are illustrated in figure 5. First, the QAIDS-IDS is strongly rejected in favor of an extended QES-IDS for this data set, for both income distribution estimates, and at both stages of the NLSURE estimation process. Second, the T3PLN density appears to fit the data substantially better than the MAXENT density, as measured by the residual sum of squares, for both iterations on the residual covariance matrix. However, the resulting estimates for the first- and second-order income coefficients, γ and δ, respectively, as well as the optimal values for the Box-Cox parameters, κ and λ, are statistically very similar in the two specifications for the income distribution. Third, the residual sum of squares appears to have a discontinuity in the interval $\kappa \in [.05, .10]$, which appears even with increments in κ of .001 in that interval and with the starting values for the other parameters in the nonlinear estimation procedure initialized at each fixed value for κ in the range 0.001 to 0.200. However, this discontinuity decreases substantially in the second NLSURE step.

Table 3 presents the individual equation summary statistics for the T3PLN model of income distribution and table 4 presents this information for the MAXENT model. As would be expected from the differences in the residual sums of squares, the R^2 measure of fit for the former tend to be slightly higher than for the latter in most of the demand equations, although this is reversed for nine out of the twenty-one goods. On the other hand, three of the Durbin-Watson statistics for serial correlation (denoted by D-W in the tables) are considerably closer to 2.0 in the T3PLN model relative to the MAXENT model (cheese, fish, and process fruit), although only the first of these (i.e., for cheese) reverses the inference from rejecting the null hypothesis of no autocorrelation based on the lower bound for the five percent critical value of the test statistic. On the other hand, four Durbin-Watson statistics in the MAXENT model are considerably closer to 2.0 than in the T3PLN model, and none of these reverse the conclusion that there is no autocorrelation at the five percent significance level. Evidently, both versions show heteroskedasticity for a substantial number of the demands equations (seven in the T3PLN model and ten in the MAXENT model), based on the Breush-Pagan Lagrange multiplier test. Although not very surprising given the long and frequently volatile period covered by the sample set, this issue warrants further consideration.

Table 5 presents the Box-Cox price coefficient and the first- and second-order income coefficients for both versions of the model. The standard errors reported in this table are conditional on the estimate of κ due to the generated income variables nature of the demand model's parameter estimates. This, combined with the evidence of potentially serious heteroskedasticity presented in the previous two tables, implies that these standard errors should be interpreted with caution. However, in the absence of heteroskedasticity, it is possible to calculate consistent test statistics for the rank of the demand model using a Wald test. For the T3PLN version, we obtain the following:

$H_0: \gamma = \mathbf{0}$ $\quad H_1: \gamma \neq \mathbf{0}$ $\quad \chi^2(21) = 114.89$

$H_0: \delta = \mathbf{0}$ $\quad H_1: \delta \neq \mathbf{0}$ $\quad \chi^2(21) = 59.99$

$H_0: \gamma = \delta = \mathbf{0}$ $\quad H_1: \gamma \neq \mathbf{0}$ or $\delta \neq 0$ $\quad \chi^2(42) = 349.18$

Similarly, for the MAXENT version, we obtain:

$H_0: \gamma = \mathbf{0}$ $\quad H_1: \gamma \neq \mathbf{0}$ $\quad \chi^2(21) = 148.77$

$H_0: \delta = \mathbf{0}$ $\quad H_1: \delta \neq \mathbf{0}$ $\quad \chi^2(21) = 105.16$

$H_0: \gamma = \delta = \mathbf{0}$ $\quad H_1: \gamma \neq \mathbf{0}$ or $\delta \neq 0$ $\quad \chi^2(42) = 373.50$

In both cases, we are lead to reject all three versions of the null hypothesis at any standard level of significance, and therefore conclude that the full rank three QES-IDS model is a significant improvement over all of the more restrictive versions. We also conclude that any version of integrable AIDS model is significantly inferior to the corresponding alternative with the Box-Cox income parameter statistically very close to unity.

5. Conclusions

This paper presents a method to nest, test and estimate both the rank and functional form of the income terms in an incomplete system of aggregable and integrable demand equations is derived. The Maximum entropy procedure is applied to the problem of inferring the U.S. income distribution using annual time series data on quintile and top five percentile income ranges and intra-quintile and top five percentile mean incomes. The results obtained with the MAXENT income distribution are compared and contrasted to those obtained with a parametric truncated three-parameter lognormal income distribution. The estimates for the year-to-year income distribution are combined with annual time series data on the U.S. consumption of and retail prices for twenty-one food items over the period 1919–95, excluding 1942–46 to account for the structural impacts of World War II.

The empirical results suggest that all integrable versions of the AIDS model are strongly rejected by this data set, in favor of a full rank three extended QES-IDS. This has potentially significant implications for future demand analysis, particularly with respect to food consumption using aggregate market-level data sets. For example, in his model of the demand for dairy products, Agnew (1998) finds the nonhomothetic, integrable rank two AIDS model to be substantially responsible for rejections of the implications of consumer choice theory – both symmetry and curvature – as well as a similar result as is reported here regarding the inferiority in all statistical respects relative to an extended LES model specification. The extreme level of confidence with which we reject the AIDS forms here suggests that a similar finding is likely. This, of course, must be left for future research.

The empirical results presented in this paper regarding the demand for U.S. food consumption are somewhat limited in their scope and interpretation. The primary reason for this is the fact that all other parameter estimates are conditional on the estimated Box-Cox parameter for the income coefficient. On the other hand, however, if we were to assume *a priori* that a QES model is the best specification – which of course at this stage of the game is unfair play – then we could interpret the remaining parameter estimates in the usual manner. It is interesting to note that, given the QES specification, the moments required from the income distribution for exact aggregation are precisely the mean and the variance. This is an interesting implication of the present study in its own right. A second limitation of the current paper's empirical results is the presence of a significant level of heteroskedasticity. This property seriously impacts the inferences that are possible regarding the unknown parameters and model structure, and therefore warrants further consideration. Finally, no attempt is made in the present empirical work to test or impose the appropriate curvature restrictions necessary for the demand model to be logically consistent with *weak integrability*, and therefore the maximization hypothesis. Consequently, the empirical results reported here cannot be use for welfare analysis.

References

Agnew, G. K. *Linquad* Unpublished M.S. Thesis, Department of Agricultural and Resource Economics, University of Arizona, Tucson, 1998.

Anderson, G. and R. Blundell. "Testing Restrictions in a Flexible Dynamic Demand System: An Application to Consumer's expenditure in Canada." *Review of Economic Studies* 50 (1983): 397-410.

Blackorby, C. and T. Shorrocks, Eds. *Separability and Aggregation: collected Works of W. M. Gorman, Volume I* Oxford: Clarendon Press, 1995.

Browning, M. and C. Meghir. "The Effects of Male and Female Labor Supply on Commodity Demands." *Econometrica* 59 (1991): 925-951.

Buse, A. "Testing Homogeneity in the Linearized Almost Ideal Demand System." *American Journal of Agricultural Economics* 80 (1998): 208-220.

Csiszár, I. "Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems." *The Annals of Statistics* 19 (1991): 2032-2066.

Deaton, A. and Muellbauer, J. "An Almost Ideal Demand system." *American Economic Review*, 70 (1980): 312-326.

Dhrymes, P. *Mathematics for Econometricians.* Springer-Verlag, New York: 1984.

Diewert, E. "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function." *Journal of Political Economy* 71 (1971): 481-507.

Epstein, L. "Integrability of Incomplete Systems of Demand Functions." *Review of Economic Studies* 49 (1982): 411-425.

Golan, A., G. Judge, and D. Miller *Maximum Entropy Econometrics: Robust Estimation with Limited Data,* New York: John Wiley and Sons, 1996.

Gokhale, D. V. and S. Kullback. *The Information in Contingency Tables,* New York: Marcel Dekker, 1978.

Gorman, W. M. "Consumer Budgets and Price Indices." Unpublished typescript, 1965. Published as Chapter 5 in Blackorby, C. and T. Shorrocks, Eds. *Separability and Aggregation: collected Works of W. M. Gorman, Volume I* Oxford: Clarendon Press, 1995: 61-88.

_________. "Some Engel Curves." In A. Deaton, ed. *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone,* Cambridge: Cambridge University Press, 1981. Republished as Chapter 20 in Blackorby, C. and T. Shorrocks, Eds. *Separability and Aggregation: collected Works of W. M. Gorman, Volume I* Oxford: Clarendon Press, 1995: 351-376.

Howe, H., R. A. Pollak, and T. J. Wales. "Theory and Time Series Estimation of the Quadratic Expenditure System." *Econometrica* 47 (1979): 1231-1247.

Jaynes, E. "Information Theory and Statistical Mechanics." *Physics Review* 106 (1957a): 620-30.

__________ "Information Theory and Statistical Mechanics II." *Physics Review* 108 (1957b): 171-90.

__________. 'Prior Information and Ambiguity in Inverse Problems." in D. McLaughlin, ed., *Inverse Problems,* Providence RI: American Mathematical Society, 1984: 151-166.

Johnson, N. L., S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Volume I*, 2nd Edition, New York: John Wiley and Sons, 1994.

Kullback, J. *Information Theory and Statistics,* New York: John Wiley and Sons, 1959.

Kullback, J. and R. Leibler. "On Information and Sufficiency." *Annals of Mathematical Statistics* 4 (1951): 99-111.

LaFrance, J. T. "Linear Demand Functions in Theory and Practice." *Journal of Economic Theory*, 37, 1 (October 1985): 147-166.

_________. "Incomplete Demand Systems and Semilogarithmic Demand Models." *Australian Journal of Agricultural Economics* 34 (1990): 118-131.

__________. "An Econometric Model of the Demand for Food and Nutrition." *The 1999 Havlicek Memorial Lecture in Applied Economics*, Department of Agricultural, Environmental and Development Economics, Ohio State University, Columbus OH, 1999a.

__________. "U.S. Food and Nutrient Demand and the Effects of Agricultural Policies." *Proceedings of the Food and Agricultural Consortium*, Washington DC: 1999b.

__________. "Inferring the Nutrient Content of Food with Prior Information." *American Journal of Agricultural* Economics, 81 (1999c): 728-734.

LaFrance, J. and M. Hanemann. "The Dual Structure of Incomplete Demand Systems." *American Journal of Agricultural Economics* 71 (1989): 262-274.

Lewbell, A. "Characterizing Some Gorman Engel Curves." *Econometrica* 55 (1987): 1451-1459.

_________. "Full Rank Demand Systems." *International Economic Review* 31 (1990): 289-300.

Moschini, G. "Units of Measurement and the Stone Index in Demand System Estimation." *American Journal of Agricultural Economics* 77 (1995): 63-68.

Moschini, G. and K. Meilke. "Modeling the Pattern of Structural Change in U.S. Meat Demand." *American Journal of Agricultural Economics* 71 (1989): 253-261.

Muellbauer, J. "Aggregation, Income Distribution and Consumer Demand." *Review of Economic Studies* 42 (1975): 525-543.

________. "Community Preferences and the Representative Consumer." *Econometrica* 44 (1976): 979-999.

Pashardes, P. "Bias in Estimating the Almost Ideal Demand system with the Stone Index Approximation." *Economic Journal* 103 (1993): 908-915.

Shannon, C. "A Mathematical theory of Communication." *Bell System Technical Journal* 27 (1948): 379-423.

Tobias, J. and A. Zellner. "Further Results on the Maximum entropy Analysis of the

Multiple Regression Model." H.G.B. Alexander Research Foundation, University of Chicago. Presented at the Econometric Society meeting, June 1997.

U.S. Department of Commerce, Bureau of the Census, *Historical Statistics of the United States: Colonial Times to 1970*, Part 1, ..., Washington DC: 1972.

van Daal, J. and A. H. Q. M. Merkies. "A Note on the Quadratic Expenditure Model." *Econometrica* 57 (1989): 1439-1443.

Zellner, A. "Optimal Information Processing and Bayes Theorem." *American Statistician* (1988): 278-84.

__________. "The Maximum entropy (MAXENT): Theory and Applications." In T. Fomby and R. Hill, Eds., *Advances in Econometrics* 12 (1997): 85-105.

Zellner, A., J. Tobias, and K. Ryu. "Maximum entropy (MAXENT) Analysis of Parametric and Semiparametric Regression Models." Alexander Research Foundation, University of Chicago, 1997.

Appendix: Proofs of Lemmas

Lemma 1: It can be show that $|\boldsymbol{I}+\gamma\boldsymbol{x}'| = 1+\gamma'\boldsymbol{x}$, so that $\boldsymbol{I}+\gamma\boldsymbol{x}'$ is nonsingular and has an inverse defined by $\boldsymbol{I} - \gamma\boldsymbol{x}'/(1+\gamma'\boldsymbol{x})$ if and only if $\gamma'\boldsymbol{x} \neq -1$ (Dhrymes, 1984: 38-39). Typically, $\boldsymbol{p}$ is a vector of price indices each normalized to unity in a base period, so that $\boldsymbol{x}$ vanishes in the base period. In addition, the elements of γ, which measure the departure from homotheticity of the individual demand equations, often are quite small (e.g., the empirical results obtained by Deaton and Muellbauer). Moreover, 0º homogeneity requires that the elements of γ sum to zero. Therefore, in all cases where $\boldsymbol{x} = \theta\iota$ with $\iota = [1\ 1 \cdots 1]'$ for some $\theta \in \oplus$ (including $\theta = 0$) at a base point for the data, the matrix $\boldsymbol{I}+\gamma\boldsymbol{x}'$ is nonsingular in a neighborhood of that point. We therefore have the following property.

A. $$1+\gamma'\boldsymbol{x} \neq 0 \ \forall \ \boldsymbol{x} \in \text{q} \quad \subset \mathbb{R}^n,$$

where θ is open, has a nonempty interior, contains the line passing through **0** and ι, and includes all of $\mathbb{R}^n$ except an n-1 dimensional hyperplane with Lebesgue measure zero.

Property **A** permits us to write the LA-AIDS as a system of linear partial differential equations,

$$\frac{\partial y(\boldsymbol{x},u)}{\partial \boldsymbol{x}} - \gamma\frac{y(\boldsymbol{x},u)}{(1+\gamma'\boldsymbol{x})} = \left[\boldsymbol{I} - \frac{\gamma\boldsymbol{x}'}{(1+\gamma'\boldsymbol{x})}\right](\alpha+\boldsymbol{B}\boldsymbol{x}),$$

where use has been made of $[\boldsymbol{I} - \gamma\boldsymbol{x}'/(1+\gamma'\boldsymbol{x})]\gamma \equiv \gamma/(1+\gamma'\boldsymbol{x})$. Then, by simply noting that

$$\frac{\partial}{\partial \boldsymbol{x}}\left[\frac{y(\boldsymbol{x},u)}{1+\gamma'\boldsymbol{x}}\right] = \left[\frac{\partial y(\boldsymbol{x},u)}{\partial \boldsymbol{x}} - \gamma\frac{y(\boldsymbol{x},u)}{(1+\gamma'\boldsymbol{x})}\right]\frac{1}{(1+\gamma'\boldsymbol{x})},$$

we can multiply by $1/(1+\gamma'\boldsymbol{x})$ to make the left-hand-side an exact differential. Consequently, Slutsky symmetry is equivalent to symmetry of the $n\times n$ matrix

$$\frac{\partial}{\partial \boldsymbol{x}'}\left\{\frac{1}{(1+\gamma'\boldsymbol{x})}\left[\boldsymbol{I} - \frac{\gamma\boldsymbol{x}'}{(1+\gamma'\boldsymbol{x})}\right](\alpha+\boldsymbol{B}\boldsymbol{x})\right\} =$$

$$\frac{\boldsymbol{B}}{(1+\gamma'\boldsymbol{x})} - \frac{\left[(\alpha+\boldsymbol{Bx})\gamma' + \gamma(\alpha' + \boldsymbol{x}'\boldsymbol{B}' + \boldsymbol{x}'\boldsymbol{B})\right]}{(1+\gamma'\boldsymbol{x})^2} - \frac{2(\alpha'\boldsymbol{x} + \boldsymbol{x}'\boldsymbol{Bx})\gamma\gamma'}{(1+\gamma'\boldsymbol{x})^3}.$$

Imposing symmetry on each of the terms associated with like powers of $(1 + \gamma'\boldsymbol{x})$ and ignoring terms that are automatically symmetric, we obtain $\boldsymbol{B} = \boldsymbol{B}'$ and $\gamma\boldsymbol{x}'\boldsymbol{B} = \boldsymbol{B}'\boldsymbol{x}\gamma'$. There are two ways that these conditions are satisfied simultaneously $\forall\ \boldsymbol{x} \in \theta$: (i) $\gamma \neq \boldsymbol{0}$ and $\boldsymbol{B} = \beta_0\gamma\gamma'$ for some $\beta_0 \in \oplus$ (including $\beta_0 = 0$); and (ii) $\gamma = \boldsymbol{0}$ and $\boldsymbol{B} = \boldsymbol{B}'$.

Case (i) gives the LA-AIDS model in the form

$$\frac{\partial y}{\partial \boldsymbol{x}} = \alpha + \beta_0\gamma\gamma'\boldsymbol{x} + \gamma\left(\frac{y - \alpha'\boldsymbol{x} - \beta_0(\gamma'\boldsymbol{x})^2}{1+\gamma'\boldsymbol{x}}\right) = \alpha + \gamma\left(\frac{y - \alpha'\boldsymbol{x} + \beta_0\gamma'\boldsymbol{x}}{1+\gamma'\boldsymbol{x}}\right).$$

This is a very simple system of linear first-order partial differential equations. Noting that

$$\frac{\partial}{\partial \boldsymbol{x}}\left(\frac{\alpha'\boldsymbol{x}}{1+\gamma'\boldsymbol{x}}\right) = \left[\alpha - \gamma\left(\frac{\alpha'\boldsymbol{x}}{1+\gamma'\boldsymbol{x}}\right)\right]\frac{1}{(1+\gamma'\boldsymbol{x})},$$

and that

$$\frac{\partial}{\partial \boldsymbol{x}}\left\{\beta_0\left[\ln(1+\gamma'\boldsymbol{x}) - \left(\frac{\gamma'\boldsymbol{x}}{1+\gamma'\boldsymbol{x}}\right)\right]\right\} = \frac{\beta_0\gamma\gamma'\boldsymbol{x}}{(1+\gamma'\boldsymbol{x})^2},$$

combining these two equations with equation (4), and integrating with respect to $\boldsymbol{x}$, we obtain the logarithmic expenditure function as

$$y(\boldsymbol{x},u) = \alpha'\boldsymbol{x} + \beta_0\left[(1+\gamma'\boldsymbol{x})\ln(1+\gamma'\boldsymbol{x}) - \frac{\gamma'\boldsymbol{x}}{(1+\gamma'\boldsymbol{x})}\right] + (1+\gamma'\boldsymbol{x})u,$$

with an obvious normalization for the utility index.

Case (ii) generates the homothetic LA-AIDS demand model,

$$\frac{\partial y}{\partial \boldsymbol{x}} = \alpha + \boldsymbol{Bx},$$

which gives the logarithmic expenditure function as

$$y(\boldsymbol{x},u) = \alpha'\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x} + u\,,$$

again with an obvious normalization.

This establishes necessity of the parameter restrictions for integrability. On the other hand, $\gamma = \mathbf{0}$ trivially gives an LA-AIDS form. To show sufficiency when $\gamma \neq \mathbf{0}$, write

$$y - \boldsymbol{x}'\frac{\partial y}{\partial \boldsymbol{x}} = y - \boldsymbol{x}'\left[\alpha + \beta_0\gamma\gamma'\boldsymbol{x} + \gamma\left(\frac{y - \alpha'\boldsymbol{x} - \beta_0(\gamma'\boldsymbol{x})^2}{1+\gamma'\boldsymbol{x}}\right)\right] = \frac{y - \alpha'\boldsymbol{x} - \beta_0(\gamma'\boldsymbol{x})^2}{1+\gamma'\boldsymbol{x}}.$$

Direct substitution gives the equivalent LA-AIDS form,

$$\frac{\partial y}{\partial \boldsymbol{x}} = \alpha + \beta_0\gamma\gamma'\boldsymbol{x} + \gamma\left(y - \boldsymbol{x}'\frac{\partial y}{\partial \boldsymbol{x}}\right). \qquad \blacksquare$$

Lemma 2: Symmetry of the Slutsky substitution terms is equivalent to symmetry of the $n \times n$ matrix with typical element

$$s_{ij} = \beta_{ij} + \sum_{k=1}^{n}\theta_{ijk}x_k + \gamma_i\left[\alpha_j + \sum_{k=1}^{n}\beta_{jk}x_k + \tfrac{1}{2}\sum_{k=1}^{n}\sum_{\ell=1}^{n}\theta_{jk\ell}x_k x_\ell + \gamma_j y\right].$$

To show necessity, we will derive the implications of symmetry, $s_{ij} = s_{ji}\ \forall\ i,j$. These implications can be conveniently grouped into three sets:

(a) $$\beta_{ij} + \gamma_i\alpha_j = \beta_{ji} + \gamma_j\alpha_i\,;$$

(b) $$\sum_{k=1}^{n}(\theta_{ijk} + \gamma_i\beta_{jk})x_k = \sum_{k=1}^{n}(\theta_{jik} + \gamma_j\beta_{ik})x_k\,;$$

and

(c) $$\gamma_i\sum_{k=1}^{n}\sum_{l=1}^{n}\theta_{jkl}x_k x_l = \gamma_j\sum_{k=1}^{n}\sum_{l=1}^{n}\theta_{ikl}x_k x_l$$

From (a), it follows that $\alpha \equiv \hat{\alpha} - \gamma\hat{\alpha}_0$, where $\hat{\alpha}_i = (\beta_{i1} - \beta_{1i})/\gamma_1$ and $\hat{\alpha}_0 = -\alpha_1/\gamma_1$. Substituting the right-hand-side for each α_i back into (a) implies $B \equiv [\beta_{ij} + \gamma_i\hat{\alpha}_j]$ is symmetric, equivalently $\left[\beta_{ij}\right]=\boldsymbol{B}-\gamma\alpha'$ for some symmetric matrix $\boldsymbol{B}$.

Now turning to (b), we use a result of LaFrance and Hanemann (1989, theorem 2, p. 266) to obtain $\theta_{ijk}+\gamma_i\beta_{jk}=\theta_{jik}+\gamma_j\beta_{ik}\ \forall\ i,j,k$, which we will return to in a moment. First, however, we apply the same result of LaFrance and Hanemann to (c) to get $\gamma_i\theta_{jkl}=\gamma_j\theta_{ikl}\ \forall\ i,j,k,l$, which in turn implies that, for each *i*, the $n\times n$ matrix $[\theta_{ikl}]=\gamma_i\boldsymbol{C}$ where $\boldsymbol{C}$ is a symmetric matrix with typical element $c_{kl}=\theta_{1kl}/\gamma_1$. Combining this with (b) gives $\gamma_i(c_{jk}+b_{jk})=\gamma_j(c_{ik}+b_{ik})\ \forall\ i,j,k$. Exploiting $\gamma_1\neq 0$ and the symmetry of both $\boldsymbol{B}$ and $\boldsymbol{C}$ then give $(b_{ij}+c_{ij})=(b_{11}+c_{11})\gamma_i\gamma_j/\gamma_1^2$, so that $\boldsymbol{B}$ and $\boldsymbol{C}$ are related by $\boldsymbol{C}=-(\boldsymbol{B}+\varepsilon\gamma\gamma')$, where $\varepsilon=-(b_{11}+c_{11})/\gamma_1^2$.

Combining all of these implications, the transformed demands can be written in matrix notation as

$$\frac{\partial y(\boldsymbol{x},u)}{\partial \boldsymbol{x}}=\hat{\alpha}-\hat{\alpha}_0\gamma+(\boldsymbol{B}-\gamma\hat{\alpha}')\boldsymbol{x}-\tfrac{1}{2}\gamma\boldsymbol{x}'(\boldsymbol{B}+\varepsilon\gamma\gamma')\boldsymbol{x}+\gamma y(\boldsymbol{x},u)$$

$$=\hat{\alpha}+\boldsymbol{Bx}+\gamma\left[y(\boldsymbol{x},u)-\hat{\alpha}_0-\hat{\alpha}'\boldsymbol{x}-\tfrac{1}{2}\boldsymbol{x}'(\boldsymbol{B}+\varepsilon\gamma\gamma')\boldsymbol{x}\right].$$

Now, note that adding and subtracting $\varepsilon\gamma$ and $\varepsilon\gamma\gamma'\boldsymbol{x}$ has no affect on the transformed demands. Therefore, let $\tilde{\alpha}=\hat{\alpha}+\varepsilon\gamma$, $\tilde{\boldsymbol{B}}=\boldsymbol{B}+\varepsilon\gamma\gamma'$, and $\alpha_0=\hat{\alpha}_0+\varepsilon$, and rewrite the partial differential equations in the equivalent form

$$\frac{\partial y(\boldsymbol{x},u)}{\partial \boldsymbol{x}}=\tilde{\alpha}+\tilde{\boldsymbol{B}}\boldsymbol{x}+\gamma\left[y(\boldsymbol{x},u)-\alpha_0-\tilde{\alpha}'\boldsymbol{x}-\tfrac{1}{2}\boldsymbol{x}'\tilde{\boldsymbol{B}}\boldsymbol{x}\right].$$

Finally, the integrating factor $e^{-\gamma'\boldsymbol{x}}$ makes the partial differential equations exact,

$$\frac{\partial}{\partial \boldsymbol{x}}\left[y(\boldsymbol{x},u)e^{-\gamma'\boldsymbol{x}}\right]=\left[\tilde{\alpha}+\tilde{\boldsymbol{B}}\boldsymbol{x}-\gamma\left(\alpha_0+\tilde{\alpha}'\boldsymbol{x}+\tfrac{1}{2}\boldsymbol{x}'\tilde{\boldsymbol{B}}\boldsymbol{x}\right)\right]e^{-\gamma'\boldsymbol{x}}$$

$$=\frac{\partial}{\partial \boldsymbol{x}}\left[\left(\alpha_0+\tilde{\alpha}'\boldsymbol{x}+\tfrac{1}{2}\boldsymbol{x}'\tilde{\boldsymbol{B}}\boldsymbol{x}\right)e^{-\gamma'\boldsymbol{x}}\right],$$

with complete solution class given by

$$y(\boldsymbol{x},u) = \alpha_0 + \tilde{\alpha}'\boldsymbol{x} + \tfrac{1}{2}\boldsymbol{x}'\tilde{\boldsymbol{B}}\boldsymbol{x} + \delta(\tilde{\boldsymbol{p}},u)\,\mathrm{e}^{\gamma'\boldsymbol{x}}.$$

Sufficiency is demonstrated by applying Hotelling's/Shephard's lemma. ■

Lemma 3. We begin by proving the lemma for simple polynomials in *m*. Dropping the ***p*** arguments for notational clarity, the Slutsky matrix can be written as

$$\boldsymbol{S} = \sum_{i=0}^{k} \frac{\partial \alpha_i}{\partial \boldsymbol{p}'} m^i + \sum_{i=1}^{k}\sum_{j=0}^{k} i\alpha_i\alpha'_j m^{i+j-1},$$

where each $\partial\alpha_i/\partial\boldsymbol{p}'$ is an $n \times n$ matrix. By the continuity of each term in ***S***, symmetry requires that each like power of *m* must have a symmetric coefficient matrix. All such matrices In reply to: powers of *m* from *k*+1 through 2*k*-2 involve nontrivial symmetry conditions without any $\partial\alpha_i/\partial\boldsymbol{p}'$ terms (The $m^{2k\text{-}1}$ matrix only involves $\alpha_k\alpha'_k$, which is clearly symmetric.) We will combine terms in the like powers of *m* and work backwards recursively from the $m^{2k\text{-}2}$ matrix; then

$$(k-1)\alpha_{k-1}\alpha'_k + k\alpha_k\alpha'_{k-1}$$

is symmetric if and only if $\alpha_{k-1} \equiv \varphi_{k-1}\alpha_k$, say, for some $\varphi_{k-1} : \mathbb{R}^n_+ \to \mathbb{R}$. Similarly,

$$(k-2)\alpha_{k-2}\alpha'_k + (k-1)\alpha_{k-1}\alpha'_{k-1} + k\alpha_k\alpha'_{k-2}$$

is symmetric if and only if $\alpha_{k-2} \equiv \varphi_{k-2}\alpha_k$ for some $\varphi_{k-2} : \mathbb{R}^n_+ \to \mathbb{R}$. Now, to see how the recursion applies, consider the $m^{2k\text{-}4}$ matrix for $k \geq 3$,

$$(k-3)\alpha_{k-3}\alpha'_k + (k-2)\alpha_{k-2}\alpha'_{k-1} + (k-1)\alpha_{k-1}\alpha'_{k-2} + k\alpha_k\alpha'_{k-3}.$$

Note that symmetry of the middle two terms follows from the two previous results, since then we have $\alpha_{k-2}\alpha'_{k-1} = \varphi_{k-2}\varphi_{k-1}\alpha_k\alpha'_k$ and $\alpha_{k-1}\alpha'_{k-2} = \varphi_{k-1}\varphi_{k-2}\alpha_k\alpha'_k$. Also, note that

$$(k-3)\left(\alpha_{k-3}\alpha_k' + \alpha_k\alpha_{k-3}'\right)$$

is always symmetric. Therefore, the m^{2k-4} matrix is symmetric if and only if $3\alpha_k\alpha_{k-3}'$ is, which requires that $\alpha_{k-3} \equiv \varphi_{k-3}\alpha_k$ for some $\varphi_{k-3} : \mathbb{R}_+^n \to \mathbb{R}$. This completes the proof for all $k \leq 5$.

If $k > 5$, then for each j such that $4 \leq j \leq k$-1, ..., we group like terms, substitute $\alpha_{k-i} \equiv \varphi_{k-i}\alpha_k$ for each $i < j$, and exploit symmetry of the matrix $\left(\alpha_{k+1-j}\alpha_k' + \alpha_k\alpha_{k+1-j}'\right)$, which sequentially requires that each matrix of the following form is symmetric:

$$(j-1)\alpha_k\alpha_{k+1-j}' + \sum_{i=1}^{j-2}(k-i)\varphi_{k-i}\varphi_{k+1+i-j}\alpha_k\alpha_k' .$$

Each of these matrices is symmetric if, and only if, $\alpha_{k+1-j} \equiv \varphi_{k+1-j}\alpha_k$ for some $\varphi_{k+1-j} : \mathbb{R}_+^n \to \mathbb{R}$. Note that j = 4 gives us the condition for α_{k-3}, while j = k-1 gives us the condition for α_2, thereby proving inductively for any k that all of the α_i terms, i = 2, ..., k, are proportional. Therefore, the rank of the matrix $[\alpha_{ij}]$ is no greater than three.

To see that these arguments apply to polynomial expansions of the PIGL and PIGLOG models, note that

$$\frac{\partial e(\boldsymbol{p},u)^{\kappa}}{\partial \boldsymbol{p}} = \kappa \cdot e(\boldsymbol{p},u)^{\kappa-1} \cdot \left(\frac{\partial e(\boldsymbol{p},u)}{\partial \boldsymbol{p}}\right) = \alpha_0(\boldsymbol{p}) + \alpha_1(\boldsymbol{p})e(\boldsymbol{p},u)^{\kappa} + \alpha_2(\boldsymbol{p})e(\boldsymbol{p},u)^{2\kappa} + \cdots,$$

has the generalized PIGL form of Gorman and Lewbell,

$$\boldsymbol{q} = \alpha_0(\boldsymbol{p}) \cdot m^{1-\kappa} + \alpha_1(\boldsymbol{p}) \cdot m + \alpha_2(\boldsymbol{p}) \cdot m^{1+\kappa} + \cdots,$$

while

$$\frac{\partial \ln\left(e(\boldsymbol{p},u)\right)}{\partial \boldsymbol{p}} = \left(\frac{\partial e(\boldsymbol{p},u)/\partial \boldsymbol{p}}{e(\boldsymbol{p},u)}\right) = \alpha_0(\boldsymbol{p}) + \alpha_1(\boldsymbol{p})\ln\left(e(\boldsymbol{p},u)\right) + \alpha_2(\boldsymbol{p})\left[\ln\left(e(\boldsymbol{p},u)\right)\right]^2 + \cdots,$$

has the generalized PIGLOG form of Gorman and Lewbell

$$\boldsymbol{q} = \alpha_0(\boldsymbol{p}) \cdot m + \alpha_1(\boldsymbol{p}) \cdot m \cdot \ln(m) + \alpha_2(\boldsymbol{p}) \cdot m \cdot \left[\log(m)\right]^2 + \cdots .$$

Thus, the lemma applies to these models as well as to polynomials in income, and the rank three part of Gorman's theorem is a simple corollary to symmetry for this class of demand models. ■

Table 1. Prediction Equations for U.S. Income Distribution Percentile Lower Limits.

1946 – 1997

	Constant	$\ln(\mu_{pc,t})$	$[\ln(\mu_{pc,t})]^2$	Unemp	R^2	D–W
$\ln(m_{1,t})$	-5.568 (.9566)	2.514 (.2278)	-.09593 (.01332)	-1.439 (.3638)	.9990	1.621

$$e_{1,t} = \underset{(.1093)}{.6529} e_{1,t-1} + \varepsilon_{1,t}$$

1928, 1935/36, 1941 and 1946 – 1997

	Constant	$\ln(m_{i-1,t})$	$[\ln(m_{i-1,t})]^2$	$[\ln(m_{i-1,t})]^3$	R^2	D–W
$\ln(m_{2,t})$	-2.088 (1.942)	2.101 (.7209)	-.1516 (.08854)	.006813 (.003597)	.9994	1.459
$\ln(m_{3,t})$	7.242 (1.178)	1.177 (.4093)	.2223 (.0471)	-.007330 (.001794)	.9999	1.237
$\ln(m_{4,t})$	5.674 (1.151)	-.4928 (.3824)	.1331 (.04210)	-.003720 (.001535)	.9999	1.067
$\ln(m_{5,t})$	13.21 (3.505)	-2.569 (1.113)	.3267 (.1173)	-.009746 (.004095)	.9993	1.523

Numbers in parentheses are estimated least squares standard errors.

D–W is the Durbin Watson statistic for autocorrelated errors.

Table 2. Prediction Equations for U.S. Income Distribution Percentile Means.

1946 – 1997

	Constant	$\ln(\mu_{pc,t})$	$[\ln(\mu_{pc,t})]^2$	Unemp	R^2	D–W
$\ln(\mu_{1,t})$	-5.719 (2.819)	2.429 (.6691)	-.09083 (.03915)	-1.982 (.5293)	.9978	1.621

$$e_{1,t} = \underset{(.0595)}{.8943}\, e_{1,t-1} + \varepsilon_{1,t}$$

1928, 1935/36, 1941 and 1946 – 1997

	Const	$\ln(\mu_{i-1,t})$	$[\ln(\mu_{i-1,t})]^2$	$[\ln(\mu_{i-1,t})]^3$	$\ln(\mu_{pc,t})$	Unemp	R^2	D–W
$\ln(\mu_{2,t})$	7.286 (1.275)	3.190 (.5510)	-.3061 (.07212)	.01057 (.003251)	.5804 (.03359)		.9998	1.251
$\ln(\mu_{3,t})$	.9981 (.4022)	.7409 (.05770)	.008884 (.006234)		.1165 (.05418)	.3149 (.08553)	.9998	1.351
$\ln(\mu_{4,t})$	2.661 (.2268)	.5136 (.02985)	.03200 (.003515)		-.07413 (.03695)		.9999	1.056
$\ln(\mu_{5,t})$	27.92 (5.472)	-6.172 (1.804)	.6920 (.1899)	-.01980 (.006797)	.5699 (.1390)	-1.399 (.2621)	.9990	1.365
$\ln(\mu_{6,t})$	6.084 (.3858)	.2495 (.08777)	.06530 (.003334)		-.5704 (.03234)	-.5396 (.1020)	.9997	1.444

Numbers in parentheses are estimated least squares standard errors.

D–W is the Durbin Watson statistic for autocorrelated errors.

Table 3. Equation Summary Statistics for the Truncated 3-Parameter Lognormal U.S. Income Distribution

Equation	R^2	Durbin-Watson	LM Het. Test	P-Value
Fresh Milk and Cream	.9975	1.935	3.713	.054
Butter	.9971	1.559	6.839	.009
Cheese	.9977	1.353	5.349	.021
Frozen Dairy Products	.9661	1.333	.8390	.360
Canned and Powdered Milk	.9648	1.287	2.449	.118
Beef and Veal	.9741	1.280	2.367	.124
Pork	.9266	1.315	.0363	.849
Lamb, Mutton and Goat	.9569	1.455	5.835	.016
Fish	.9899	1.665	.3366	.562
Poultry	.9628	1.098	3.882	.049
Fresh Citrus Fruit	.8474	2.084	4.861	.027
Fresh Noncitrus Fruit	.9668	2.628	15.88	.000
Fresh Vegetables	.9834	1.790	.00895	.925
Potatoes and Sweet Potatoes	.9671	1.869	.2154	.643
Processed Fruit	.9869	1.829	9.921	.002
Processed Vegetables	.9785	1.554	.7384	.390
Eggs	.9747	1.771	.00394	.950
Fats and Oils, Excluding Butter	.9983	1.551	2.494	.114
Cereals and Bakery Good	.9925	1.279	3.516	.061
Sugar and Sweeteners	.9828	2.145	.3966	.529
Coffee, Tea and Cocoa	.9691	1.930	2.639	.104

Table 4. Equation Summary Statistics for the Maximum entropy U.S. Income Distribution

Equation	R^2	Durbin-Watson	LM Het. Test	P-Value
Fresh Milk and Cream	.9915	1.841	6.952	.008
Butter	.9939	1.591	13.58	.000
Cheese	.9964	1.170	1.984	.159
Frozen Dairy Products	.9748	1.323	1.443	.230
Canned and Powdered Milk	.9766	1.369	6.586	.010
Beef and Veal	.9737	1.266	3.260	.071
Pork	.9059	1.542	.0342	.853
Lamb, Mutton and Goat	.9426	1.559	5.596	.018
Fish	.9824	1.212	.5206	.471
Poultry	.9616	1.300	2.228	.136
Fresh Citrus Fruit	.9547	1.922	4.698	.030
Fresh Noncitrus Fruit	.9716	2.260	15.86	.000
Fresh Vegetables	.9949	1.749	2.5065	.113
Potatoes and Sweet Potatoes	.9620	2.217	13.90	.000
Processed Fruit	.9877	1.507	4.969	.026
Processed Vegetables	.9900	1.397	2.204	.138
Eggs	.9917	1.877	45.32	.000
Fats and Oils, Excluding Butter	.9972	1.693	11.69	.001
Cereals and Bakery Good	.9956	1.312	.2067	.649
Sugar and Sweeteners	.9735	2.091	.1889	.664
Coffee, Tea and Cocoa	.9452	1.720	.1416	.707

Table 5. Estimated Income Coefficients for the QPIGL-IDS for U.S. Food Consumption.

	Truncated 3-Parameter Lognormal: $\hat{\kappa}$ = 1.03		Bayesian Method of Moments: $\hat{\kappa}$ = 1.00	
Parameter	**Estimate**	**Conditional Standard Error**	**Estimate**	**Conditional Standard Error**
λ	.853915	.033923	.905925	.033745
γ_1	-.024642	.013093	-.024504	.015251
γ_2	$-.258821 \cdot 10^{-2}$	$.206153 \cdot 10^{-2}$	-.010249	$.140166 \cdot 10^{-2}$
γ_3	$-.911698 \cdot 10^{-3}$	$.221715 \cdot 10^{-2}$	$.984936 \cdot 10^{-3}$	$.149990 \cdot 10^{-2}$
γ_4	.017188	$.500290 \cdot 10^{-2}$	$.391228 \cdot 10^{-2}$	$.277973 \cdot 10^{-2}$
γ_5	$.301518 \cdot 10^{-2}$	$.592252 \cdot 10^{-2}$	$.573810 \cdot 10^{-3}$	$.314409 \cdot 10^{-2}$
γ_6	.022654	.010523	$.946772 \cdot 10^{-2}$	$.573816 \cdot 10^{-2}$
γ_7	.010451	$.995331 \cdot 10^{-2}$	$-.136379 \cdot 10^{-2}$	$.582051 \cdot 10^{-2}$
γ_8	$-.313399 \cdot 10^{-2}$	$.371416 \cdot 10^{-2}$	$-.245839 \cdot 10^{-2}$	$.224639 \cdot 10^{-2}$
γ_9	$.339135 \cdot 10^{-2}$	$.209848 \cdot 10^{-2}$	$.885948 \cdot 10^{-3}$	$.136181 \cdot 10^{-2}$
γ_{10}	$-.93449810^{-3}$	$.425920 \cdot 10^{-2}$	$-.270148 \cdot 10^{-3}$	$.243933 \cdot 10^{-2}$
γ_{11}	$-.968060 \cdot 10^{-3}$	$.997693 \cdot 10^{-2}$	-.020254	$.767887 \cdot 10^{-2}$
γ_{12}	.048544	.015135	.038601	.014255
γ_{13}	$.768783 \cdot 10^{-2}$	$.772512 \cdot 10^{-2}$	$.915188 \cdot 10^{-2}$	$.687591 \cdot 10^{-2}$
γ_{14}	-.020795	.020258	$-.189372 \cdot 10^{-2}$	.013863
γ_{15}	$-.550843 \cdot 10^{-2}$	$.685728 \cdot 10^{-2}$	-.012508	$.660558 \cdot 10^{-2}$
γ_{16}	.026156	$.743900 \cdot 10^{-2}$	.021646	$.662656 \cdot 10^{-2}$
γ_{17}	.011163	$.439605 \cdot 10^{-2}$	$.483915 \cdot 10^{-2}$	$.262617 \cdot 10^{-2}$
γ_{18}	$.615151 \cdot 10^{-2}$	$.353595 \cdot 10^{-2}$	$-.356145 \cdot 10^{-2}$	$.218127 \cdot 10^{-2}$
γ_{19}	.021258	.013952	$.157884 \cdot 10^{-2}$	$.929659 \cdot 10^{-2}$
γ_{20}	.019811	$.939334 \cdot 10^{-2}$	-.010722	$.601554 \cdot 10^{-2}$
γ_{21}	$.254267 \cdot 10^{-2}$	$.252764 \cdot 10^{-2}$	$.167696 \cdot 10^{-2}$	$.156889 \cdot 10^{-2}$
δ_1	$.112092 \cdot 10^{-5}$	$.486111 \cdot 10^{-6}$	$.108833 \cdot 10^{-5}$	$.416836 \cdot 10^{-6}$
δ_2	$-.187811 \cdot 10^{-7}$	$.862305 \cdot 10^{-7}$	$.176424 \cdot 10^{-6}$	$.415999 \cdot 10^{-7}$
δ_3	$.866507 \cdot 10^{-7}$	$.802000 \cdot 10^{-7}$	$-.123694 \cdot 10^{-7}$	$.389894 \cdot 10^{-7}$
δ_4	$-.463795 \cdot 10^{-6}$	$.203328 \cdot 10^{-6}$	$.494122 \cdot 10^{-7}$	$.774281 \cdot 10^{-7}$
δ_5	$-.726773 \cdot 10^{-9}$	$.226065 \cdot 10^{-6}$	$.227677 \cdot 10^{-7}$	$.847906 \cdot 10^{-7}$
δ_6	$-.518846 \cdot 10^{-6}$	$.405252 \cdot 10^{-6}$	$-.121129 \cdot 10^{-6}$	$.151601 \cdot 10^{-6}$
δ_7	$-.316928 \cdot 10^{-6}$	$.401569 \cdot 10^{-6}$	$.252274 \cdot 10^{-6}$	$.157432 \cdot 10^{-6}$
δ_8	$.137460 \cdot 10^{-6}$	$.136176 \cdot 10^{-6}$	$.138579 \cdot 10^{-6}$	$.582947 \cdot 10^{-7}$
δ_9	$.720693 \cdot 10^{-8}$	$.714625 \cdot 10^{-7}$	$.553370 \cdot 10^{-8}$	$.326526 \cdot 10^{-7}$
δ_{10}	$.250661 \cdot 10^{-6}$	$.199447 \cdot 10^{-6}$	$.159695 \cdot 10^{-6}$	$.724811 \cdot 10^{-7}$
δ_{11}	$.650096 \cdot 10^{-7}$	$.391085 \cdot 10^{-6}$	$.602407 \cdot 10^{-6}$	$.187380 \cdot 10^{-6}$
δ_{12}	$-.193640 \cdot 10^{-5}$	$.670224 \cdot 10^{-6}$	$-.116682 \cdot 10^{-5}$	$.391205 \cdot 10^{-6}$
δ_{13}	$.496807 \cdot 10^{-8}$	$.294807 \cdot 10^{-6}$	$-.284328 \cdot 10^{-7}$	$.171947 \cdot 10^{-6}$
δ_{14}	$.959843 \cdot 10^{-6}$	$.731403 \cdot 10^{-6}$	$.149585 \cdot 10^{-6}$	$.354263 \cdot 10^{-6}$
δ_{15}	$.510022 \cdot 10^{-6}$	$.285776 \cdot 10^{-6}$	$.470437 \cdot 10^{-6}$	$.175568 \cdot 10^{-6}$
δ_{16}	$-.445306 \cdot 10^{-6}$	$.328907 \cdot 10^{-6}$	$-.144042 \cdot 10^{-6}$	$.183377 \cdot 10^{-6}$
δ_{17}	$-.366017 \cdot 10^{-6}$	$.187846 \cdot 10^{-6}$	$-.152148 \cdot 10^{-6}$	$.743101 \cdot 10^{-7}$
δ_{18}	$-.159808 \cdot 10^{-6}$	$.174505 \cdot 10^{-6}$	$.104303 \cdot 10^{-6}$	$.641430 \cdot 10^{-7}$
δ_{19}	$-.525896 \cdot 10^{-6}$	$.569711 \cdot 10^{-6}$	$.201483 \cdot 10^{-6}$	$.242090 \cdot 10^{-6}$
δ_{20}	$-.288873 \cdot 10^{-6}$	$.386350 \cdot 10^{-6}$	$.336199 \cdot 10^{-6}$	$.164316 \cdot 10^{-6}$
δ_{21}	$-.138823 \cdot 10^{-7}$	$.982542 \cdot 10^{-7}$	$-.441223 \cdot 10^{-8}$	$.416300 \cdot 10^{-7}$

Figure 1. U.S. Income Distribution, 1997.

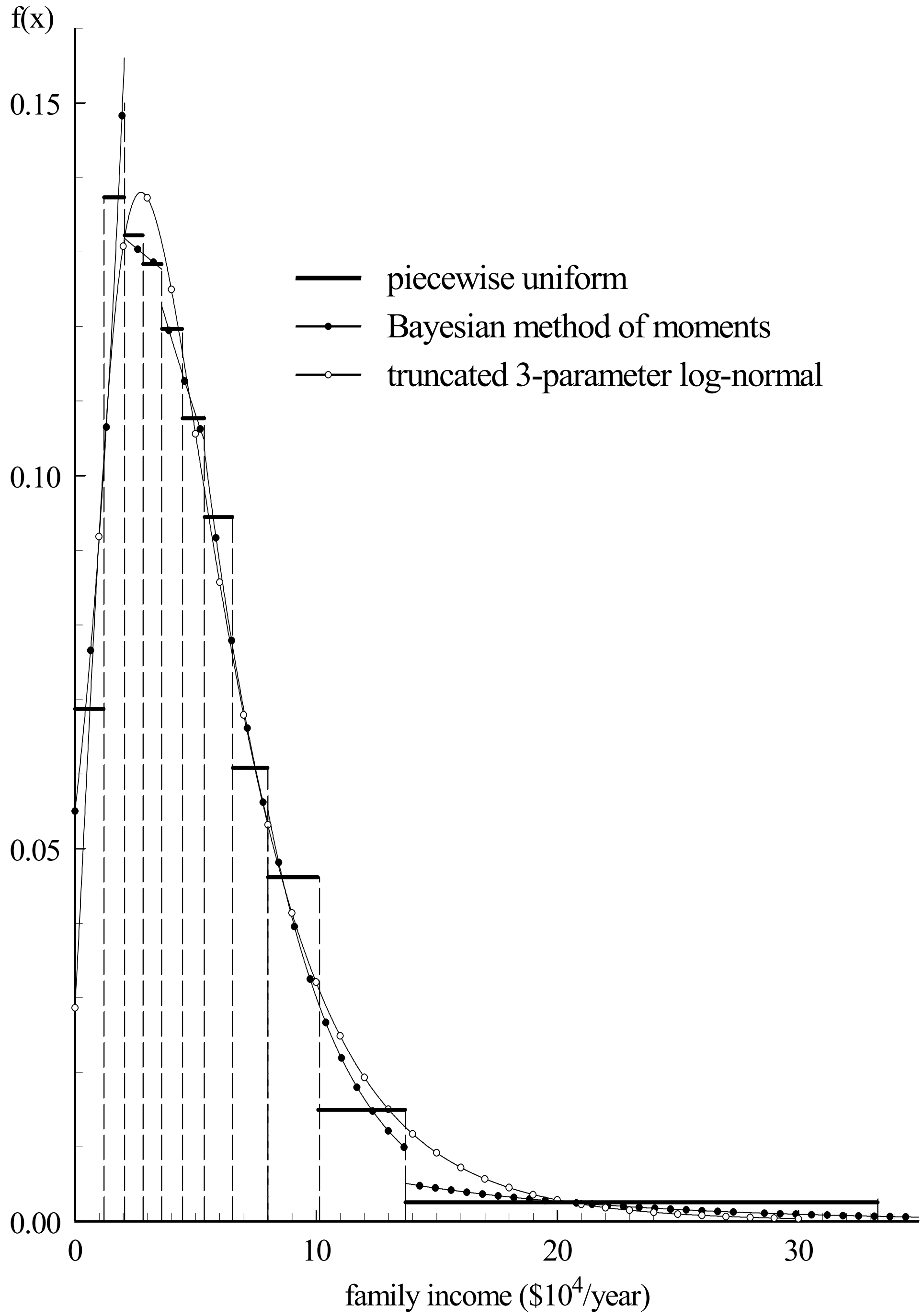

Figure 2. ln{Φ(z)/(1-Φ(z))} versus 1.6092z + .066826z³.

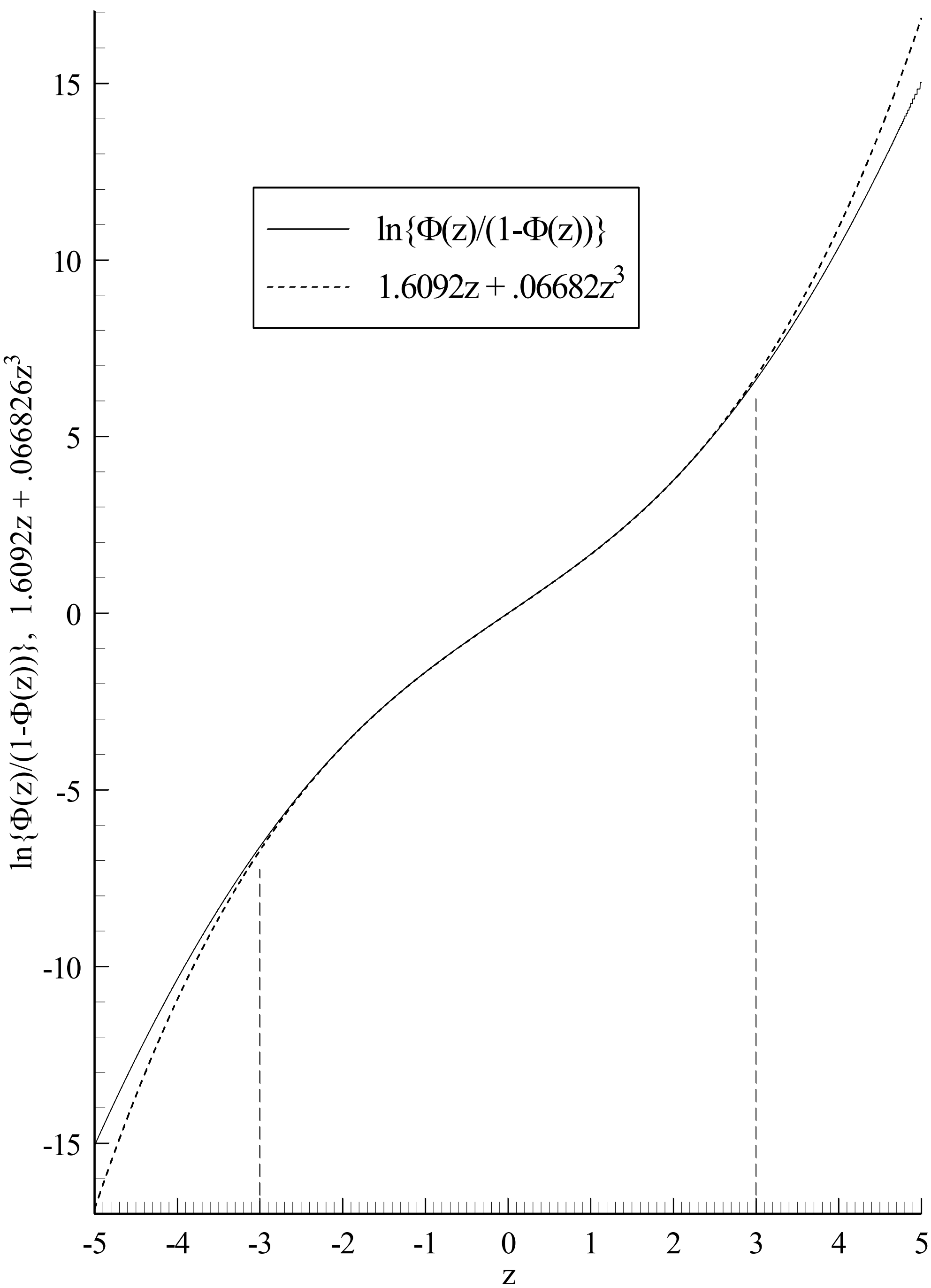

Figure 3. U.S. Income Distribution: 1910-99.

Log Quintile and Top Five Percent Income Lower Limits

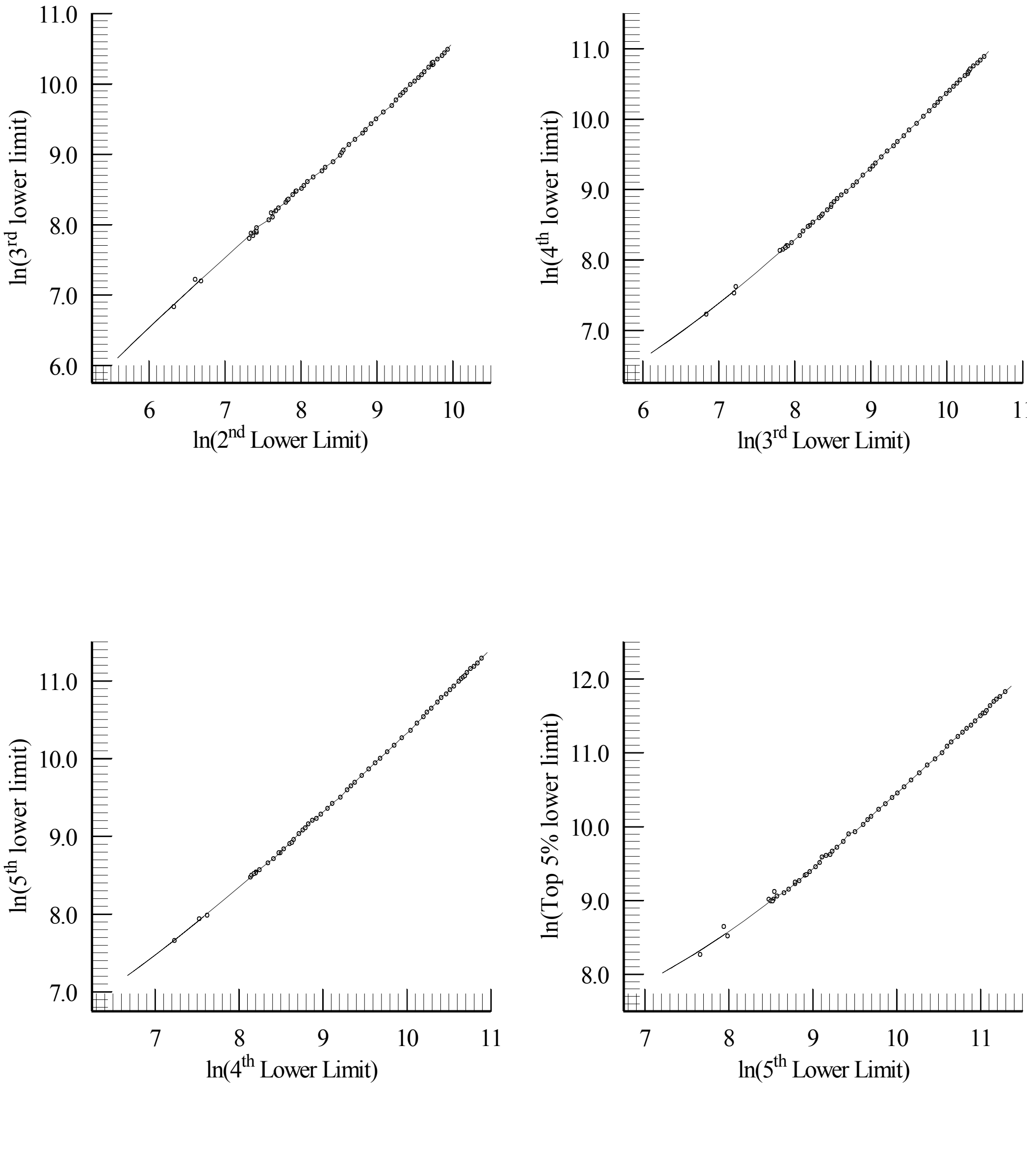

Figure 4. U.S. Income Distribution: 1910-99.

Log Quintile and Top Five Percent Income Means

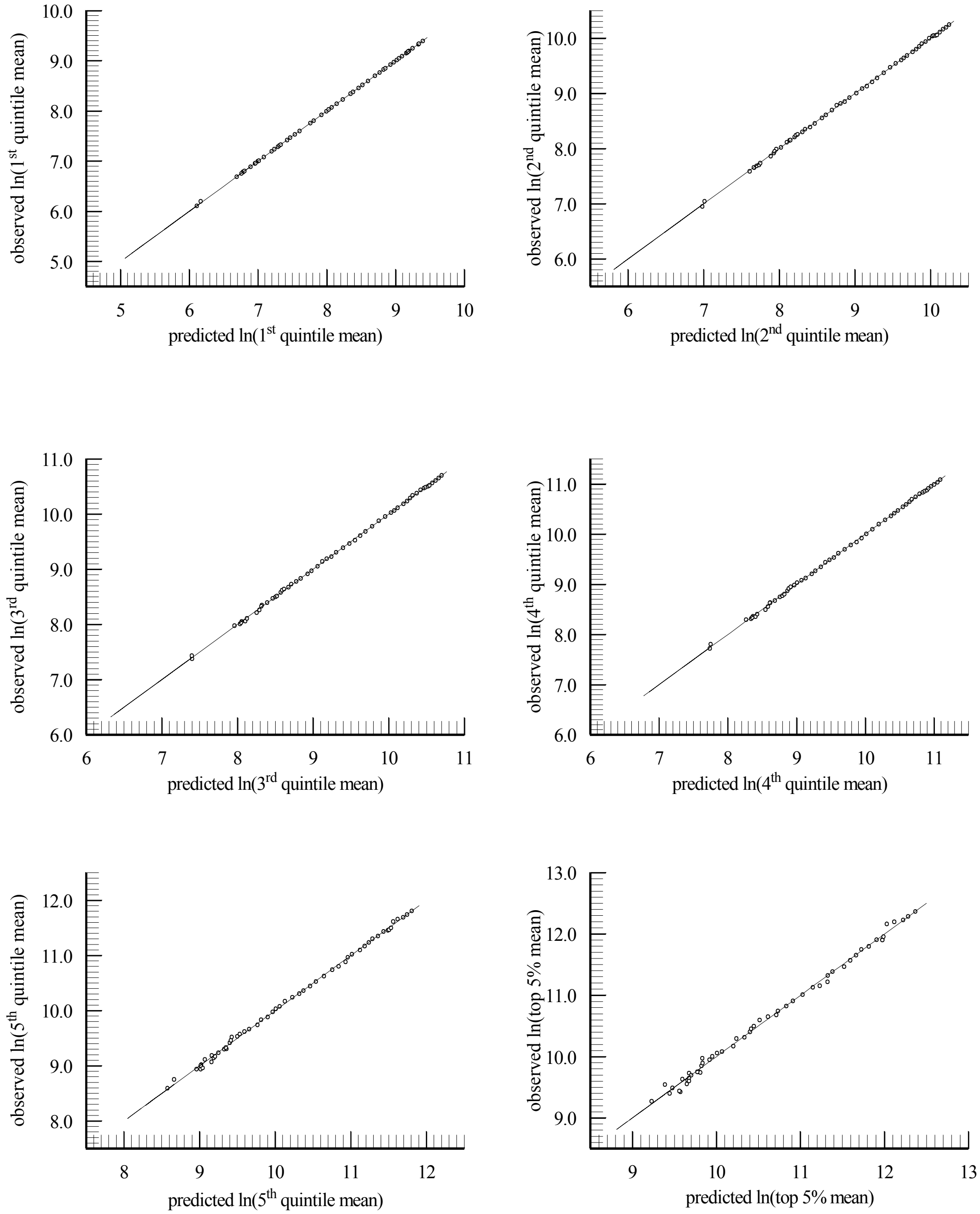

Figure 5. Nested: QPIGL: Concentrated SSR.

U.S. Food Expenditures, 1919-41 and 1947-1995

Figure 5a. First Iteration SURE Estimates

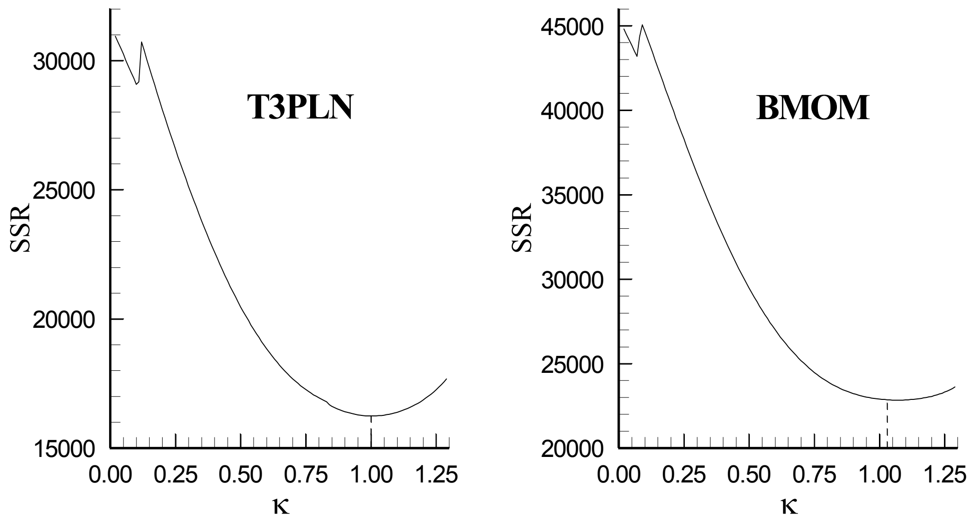

Figure 5b. Second Iteration SURE Estimates