Journal of Econometrics 107 (2002) 235–257

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Information theoretic measures of the income distribution in food demand

J.T. LaFrance[a,*], T.K.M. Beatty[a], R.D. Pope[b], G.K. Agnew[c]

[a] Department of Agricultural and Resource Economics of Economics and Giannini Foundation of Agricultural Economics, University of California, Berkeley, CA 94720-3310, USA

[b] Department of Economics, Bringham Young University, Provo, UT 84602, USA

[c] Department of Agricultral and Resource Economics, University of Arizona, Tucson, AZ 85721, USA

Abstract

A new method to nest, estimate and test the rank and functional form of the income terms in an incomplete system of demand equations is developed. Information theory is employed to infer the U.S. income distribution from data on quintile and top five percentile income ranges and intra-quintile and top five percentile mean incomes. Maximum entropy income distributions are combined with data on the U.S. demand for 21 food items to estimate U.S. food demand over the period 1919–1995, excluding 1942–1946. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: D12; E21

Keywords: Functional form; Incomplete demand systems; Information theory; Maximum entropy; Rank of demand models

1. Introduction

This paper exploits the richness of incomplete demand systems to extend aggregation in nonlinear functions of income to incomplete demand systems (hereafter IDS). We develop methods to completely nest weakly integrable versions of the almost ideal demand system (AIDS), the linear approximate AIDS (LA-AIDS), the quadratic AIDS (QAIDS), the price independent generalized linear (PIGL)—including the price independent generalized logarithmic (PIGLOG) form, the quadratic PIGL (QPIGL), an extended version of the linear expenditure system (LES)[1] and an extended quadratic

* Corresponding author. Tel.: +1-510-643-5416.

E-mail address: lafrance@are.berkeley.edu (J.T. LaFrance).

[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.

PII: S0304-4076(01)00122-1

expenditure system (QES). These methods permit us to test for and estimate the rank and functional form of the income terms in aggregable incomplete demand systems.

Rank three demand models require at least 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, $h=1,\ldots,H$, 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 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 information theory to obtain annual maximum entropy income distributions that satisfy each of the intra-range percentile and conditional mean conditions for the period 1910–1999.

These estimates of the income distribution 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 IDS specification. The results of the empirical application strongly suggest that a full rank three model is essential, and that all forms of the AIDS–IDS model are rejected in favor of a QPIGL–IDS that is nearly 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 or PIGLOG form. The third section describes the implementation of the maximum entropy procedure to estimate the U.S. income distribution. Section 4 presents a summary and discussion of the empirical results, focusing primarily on the rank of the demand model and the functional form of the income terms. The final section summarizes and concludes. Proofs of our main results and additional discussions and derivations of the modeling approach are contained in an expanded paper (LaFrance, Beatty, Pope and Agnew (2001), hereafter LBPA), which is available from the authors upon request.

2. A flexible aggregate incomplete demand system

It is a generally accepted proposition that a modern demand system should be sufficiently flexible to represent a rich set of qualitative and quantitative behaviors. Clearly,

[2] See LaFrance (1999a, b) for empirical evidence for the exclusion of World War II and the stability of U.S. food demands over this sample period. The 21 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 and bakery goods, sugar and sweeteners, and coffee, tea and cocoa.

flexibility is desired with respect to both prices and income changes. However, there is a tension that flexibility brings when applied to food or food items in the aggregate. On the one hand, the researcher might desire that demands be parsimonious in income so that demand or budget shares can be represented as simple statistics over aggregate income (e.g., Gorman, 1953, 1981). For example, homothetic or quasi-homothetic demands with constant and identical marginal propensity to consume implies that mean income is a sufficient income index to describe aggregate demands. On the other hand, as one of the oldest of empirical regularities in economics attests, the marginal propensities to consume food or food categories are not generally constant either across individuals or with respect to a person's income. Thus, Engel's law implies that more flexibility in income is desired. In this section, a flexible model of demand is presented which nests for later application many of the most common forms. This model allows one to test in the aggregate how prices and income enters food demand. The relationship of this system to Gorman's (1981) aggregable demand system is discussed. We rely heavily on previous work, particularly that of Muellbauer (1975), Deaton and Muellbauer (1980) and Howe et al. (1979), and LBPA. Throughout, the focus is on incomplete demand systems as developed in Epstein (1982), LaFrance (1985), and LaFrance and Hanemann (1989).

Over the last two decades, a number of authors have developed flexible demand systems (Lewbell, 1987, 1990; Deaton and Muellbauer, 1980; Howe et al., 1979). Two notable directions in empirical applications are the quadratic expenditure system of Howe et al. and the AIDS model of Deaton and Muellbauer. 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. [3] We begin by discussing a homothetic extension of the integrable LA-AIDS to illustrate the basic thrust of our approach. This is then extended to a rank 3 flexible system, which allows testing among IDS versions of the QAIDS and QES within a quadratic price independent generalized linear incomplete demand system (QPIGL–IDS).

2.1. Flexible incomplete extensions of LA-AIDS and AIDS

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. It proves helpful to change variables from quantities, prices, expenditures, budget shares and income to particular transformations of these variables. Let $\boldsymbol{x} \equiv \ln(\boldsymbol{p})$ and let $y(\boldsymbol{x}, u) \equiv \ln[e(\boldsymbol{p}(\boldsymbol{x}), u)]$, where $\boldsymbol{p}(\boldsymbol{x}) \equiv [e^{x_1} \cdots e^{x_n}]'$. With these definitions, it can be shown (LBPA) that one of two stems of an integrable LA-AIDS is

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

[3] Although Deaton and Muellbauer (1980, pp. 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 (e.g., Anderson and Blundell, 1983; Buse, 1998; Moschini, 1995; Moschini and Meilke, 1989; Pashardes, 1993).

This solution has exactly the same structure as the homothetic linear incomplete demand system (LIDS) of LaFrance (1985). That is, 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 $N \geqslant n+1$ goods and define Box–Cox transformations of income and prices as $m(\kappa) \equiv (m^{\kappa}-1)/\kappa$, $p_i(\lambda) \equiv (p_i^{\lambda} - 1)/\lambda$, 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

$$\boldsymbol{w} = m^{-\kappa}\boldsymbol{P}^{\lambda}[\alpha + \boldsymbol{B}\boldsymbol{p}(\lambda)], \tag{2}$$

where $\boldsymbol{P}^{\lambda} \equiv \mathrm{diag}[p_i^{\lambda}]$ is a diagonal matrix with typical diagonal element p_i^{λ}. This is shown in LBPA to be integrable with expenditure function for this PIGL–IDS

$$e(\boldsymbol{p},\tilde{\boldsymbol{p}},u) \equiv \pi(\tilde{\boldsymbol{p}})\{1 + \kappa[\alpha'\boldsymbol{p}(\lambda) + \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda) + \theta(\tilde{\boldsymbol{p}},u)]\}^{1/\kappa}, \tag{3}$$

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 can be shown that the demands associated with (2) are homothetic with income elasticities equal to $1 - \kappa$.[4]

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

$$\boldsymbol{w} = \alpha + \boldsymbol{B}\ln(\boldsymbol{p}) + \gamma[\ln(m) - \alpha_0 - \alpha'\ln(\boldsymbol{p}) - \tfrac{1}{2}\ln(\boldsymbol{p})'\boldsymbol{B}\ln(\boldsymbol{p})]. \tag{4}$$

In LBPA, integrable systems using general transformations are considered. A useful special case employs the Box–Cox transformations, $m(\kappa)$ and $\boldsymbol{p}(\lambda)$. We can write an integrable nonhomothetic PIGL–IDS that is linear in the Box–Cox income term and linear and quadratic in the Box–Cox price vector as,

$$\boldsymbol{w} = m^{-\kappa}\boldsymbol{P}^{\lambda}\{\alpha + \boldsymbol{B}\boldsymbol{p}(\lambda) + \gamma[m(\kappa) - \alpha_0 - \alpha'\boldsymbol{p}(\lambda) - \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)]\}. \tag{5}$$

The expenditure function associated with (5) is (LBPA):

$$e(\boldsymbol{p},\tilde{\boldsymbol{p}},u) \equiv \pi(\tilde{\boldsymbol{p}})\{1 + \kappa[\alpha_0 + \alpha'\boldsymbol{p}(\lambda) + \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda) + \theta(\tilde{\boldsymbol{p}},u)e^{\gamma'\boldsymbol{p}(\lambda)}]\}^{1/\kappa}, \tag{6}$$

where $\theta(\cdot)$ has the same properties as the expenditure function of Eq. (3). Unlike the homothetic case, for all (κ,λ) pairs, this flexible form in (5) and (6) allows us to

[4] For an incomplete demand system, homotheticity is defined by equality of the income elasticities of demand for all of the goods of interest. It is not necessary that any of the other demands have this same income elasticity. In particular, the common income elasticity of demand for an homothetic subset of goods is not necessarily equal to one, nor must it be constant (see LaFrance and Hanemann, 1989). This is one way in which weakly integrable incomplete demand systems are more flexible and provide a richer class of models relative to complete systems.

estimate the income aggregation function through the Box–Cox parameters. If $\kappa=\lambda=0$ we obtain the integrable AIDS model, if $\kappa=\lambda=1$ we obtain the linear–quadratic IDS (LQ-IDS) of LaFrance (1990), and for all (κ,λ) pairs we obtain an integrable PIGL–IDS.

2.2. The concept of rank, further extensions and flexibility

This nesting procedure also generalizes to demand models that include linear and quadratic terms in the Box–Cox transformation of deflated income (QES–IDS). It is useful to consider now the Gorman (1981) class of demand equations,

$$\frac{\partial y(\boldsymbol{x},u)}{\partial x}=\sum_{i=0}^{k}\alpha_i(\boldsymbol{x})g_i(y(\boldsymbol{x},u)), \tag{7}$$

where $\boldsymbol{x}$ is some transformation of prices and y is some transformation of income. 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 $h_i(m)=m^{\lambda_i}(\ln(m))^i$, 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 (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 $m^{\alpha}(\ln(m))^{\beta}$ 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.

Demand models that have *full rank* are characterized by the property that the rank of the matrix $\boldsymbol{A}(\boldsymbol{x})$ is equal to the number of its columns, that is, equal to the number of different income functions, $g_j(y)$. Clearly, complete *full rank one* demand systems must be homothetic, $\boldsymbol{q}=\alpha_0(\boldsymbol{x})m$, due to adding up. Muellbauer (1975) showed that all complete *full rank two* demand systems are either PIGL or PIGLOG, that is, either

$$\boldsymbol{q}=\alpha_0(\boldsymbol{x})m+\alpha_1(\boldsymbol{x})m^{1+\kappa}, \tag{8}$$

for some $\kappa\neq 0$, or

$$\boldsymbol{q}=\alpha_0(\boldsymbol{x})m+\alpha_1(\boldsymbol{x})m\ln(m). \tag{9}$$

Thus, the homothetic system in (2) is rank one while (4) is rank two. However, because of our particular interest in flexibility with respect to income, we now proceed to consider a rank three demand system.

Gorman (1981) showed that only three mutually exclusive cases of rank three systems are possible: (a) $m(\ln(m))^r$, where each r is an integer; (b) $m^{1+\kappa}$, where κ may or may not be an integer; and (c) $m\sin(r\ln(m))$ and $m\cos(r\ln(m))$, for some $r \geqslant 0$, 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:

$$\boldsymbol{q} = \alpha_0(\boldsymbol{x})m + \sum_{j=1}^{k} \alpha_j(\boldsymbol{x})m(\ln(m))^j; \tag{10}$$

$$\boldsymbol{q} = \alpha_0(\boldsymbol{x})m + \sum_{\kappa \in T} \beta_\tau(\boldsymbol{x})m^{1-\kappa} + \sum_{\kappa \in T} \gamma_\tau(\boldsymbol{x})m^{1+\kappa}, \tag{11}$$

where T is a set of nonzero constants; or

$$\boldsymbol{q} = \alpha_0(\boldsymbol{x})m + \sum_{\tau \in T} \beta_\tau(\boldsymbol{x})m\sin(\tau\ln(m)) + \sum_{\tau \in T} \gamma_\tau(\boldsymbol{x})m\cos(\tau\ln(m)). \tag{12}$$

where T is a set of positive constants. The case defined by (10) includes Muellbauer's (1975,1976) PIGLOG model and extensions that are polynomials in $\ln(m)$, while the case given by (11) includes polynomials in income, as well as Muellbauer's PIGL model and extensions that are polynomials in m^κ. Concerning (10) and (11), Gorman (1981) conjectured that second-order polynomials are the most general nondegenerate cases for *full rank three* demand systems. Following up on this conjecture and exploiting the methods of van Daal and Merkies (1989), Lewbell (1990) showed that, indeed, all full rank three generalizations of the PIGL and PIGLOG models are exact analogues to the QES. [5]

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 (1995), 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 form (7), it can be shown (see LBPA) that Gorman's rank theorem is a corollary to symmetry for polynomials in income, PIGL and PIGLOG functional forms.

[5] Lewbell (1990) also derived a full rank three trigonometric model of form (12), although we do not make use of that result here.

Lemma 1. *If y is m, m^{κ}, or $\ln(m)$ and the (possibly incomplete) demand system (7) is (weakly) integrable, then there exist real-valued functions, $\varphi_i : \mathbb{R}^n_+ \to \mathbb{R}$, $i=2,\ldots,k$, such that $\alpha_i(\boldsymbol{x}) \equiv \varphi_i(\boldsymbol{x})\alpha_k(\boldsymbol{x})\ \forall i \geqslant 2$.*

Thus, one cannot obtain additional income flexibility beyond that implied by a rank three IDS. Hence, we proceed with our nesting procedure by extending the rank two PIGL–IDS expenditure function in (5) 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 et al. (1979),[6]

$$\varphi(\boldsymbol{p}, m) = -\left\{ \frac{1}{[m(\kappa) - \alpha_0 - \alpha'\boldsymbol{p}(\lambda) - \frac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)]} + \delta'\boldsymbol{p}(\lambda) \right\} e^{\gamma'\boldsymbol{p}(\lambda)}. \tag{13}$$

Applying the methodology of LaFrance and Hanemann (1989), it can be shown that (13) is equivalent to an expenditure function of the form

$$e(\boldsymbol{p}, \tilde{\boldsymbol{p}}, u) \equiv \pi(\tilde{\boldsymbol{p}}) \left\{ 1 + \kappa \left[\alpha_0 + \alpha'\boldsymbol{p}(\lambda) + \frac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda) - \left(\frac{e^{\gamma'\boldsymbol{p}(\lambda)}}{\delta'\boldsymbol{p}(\lambda)e^{\gamma'\boldsymbol{p}(\lambda)} + \theta(\tilde{\boldsymbol{p}}, u)} \right) \right] \right\}^{1/\kappa}. \tag{14}$$

That is, the QPIGL–IDS expenditure function in (13) generalizes the nonhomothetic PIGL–IDS expenditure function in (6) by replacing $\theta(\tilde{\boldsymbol{p}}, u)e^{\gamma'\boldsymbol{p}(\lambda)}$ with $-[\delta'\boldsymbol{p}(\lambda) + \theta(\tilde{\boldsymbol{p}}, u)e^{-\gamma'\boldsymbol{p}(\lambda)}]^{-1}$, which produces the n-vector of parameters $\boldsymbol{\delta}$ associated with the quadratic term in supernumerary income, in addition to the n-vector of parameters $\boldsymbol{\gamma}$ associated with the linear supernumerary income term.

Finally, an application of Roy's identity to (13) generates a QPIGL–IDS in deflated expenditures form as

$$\begin{aligned} \boldsymbol{e} = m^{1-\kappa}\boldsymbol{P}^{\lambda}\{\alpha + \boldsymbol{B}\boldsymbol{p}(\lambda) + \gamma[m(\kappa) - \alpha_0 - \alpha'\boldsymbol{p}(\lambda) - \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)] \\ + [\boldsymbol{I} + \gamma\boldsymbol{p}(\lambda)']\delta[m(\kappa) - \alpha_0 - \alpha'\boldsymbol{p}(\lambda) - \tfrac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)]^2\}, \end{aligned} \tag{15}$$

where $\boldsymbol{e} = [p_1 q_1 \ldots p_n q_n]'$ is the vector of deflated expenditures on individual food items. Assuming that α and $\boldsymbol{B}$ do not completely vanish simultaneously, it follows that: (a) $\boldsymbol{\gamma} \neq \boldsymbol{0}, \boldsymbol{\delta} \neq \boldsymbol{0}$ is necessary and sufficient for a full rank three QPIGL–IDS; (b) $\boldsymbol{\gamma} \neq \boldsymbol{0}, \boldsymbol{\delta} = \boldsymbol{0}$ is necessary and sufficient for a full rank two, nonhomothetic PIGL–IDS; (c) $\boldsymbol{\gamma} = \boldsymbol{0}, \boldsymbol{\delta} \neq \boldsymbol{0}$ is necessary and sufficient for a full rank two QPIGL–IDS that excludes the linear term; and (d) $\boldsymbol{\gamma} = \boldsymbol{\delta} = \boldsymbol{0}$ is necessary and sufficient for a homothetic PIGL–IDS. Thus, we obtain a rich class of models that permits nesting, testing and estimating

[6] To see this, solve (6) for θ, transform to $\tilde{\theta} = -1/\theta$ to get $\tilde{\varphi}(\boldsymbol{p}, m) = -e^{\gamma'\boldsymbol{p}(\lambda)}/[m(\kappa) - \alpha_0 - \alpha'\boldsymbol{p}(\lambda) - \frac{1}{2}\boldsymbol{p}(\lambda)'\boldsymbol{B}\boldsymbol{p}(\lambda)]$, and then add the term $-\delta'\boldsymbol{p}(\lambda)e^{\gamma'\boldsymbol{p}(\lambda)}$ to obtain (13).

the rank and functional form of the income aggregation terms in incomplete demand systems.

3. Inferring the U.S. income distribution

We apply maximum entropy to infer the income distribution given the intra-quintile and top five percentile income ranges and conditional mean income levels.[7] In discrete form, the probability and mean constrained maximum entropy problem is to find a vector, $\boldsymbol{q}_i \geqslant \mathbf{0}$, satisfying $\iota'\boldsymbol{q}_i = \pi_i$, where $\iota = [11 \cdots 1]'$ is a vector of ones and $\boldsymbol{q}_i'\boldsymbol{m}_i = \pi_i \mu_i$ with a pre-specified vector of income levels $\boldsymbol{m}_i$ within the ith income range, usually defined as the midpoints of N equally spaced subintervals of $[\ell_{i-1}, \ell_i]$,

$$\boldsymbol{m}_i = [\, \ell_{i-1} + \tfrac{1}{2N}\Delta\ell_i \;\; \ell_{i-1} + \tfrac{3}{2N}\Delta\ell_i \;\; \cdots \;\; \ell_{i-1} + \tfrac{2N-1}{2N}\Delta\ell_i \,]', \tag{16}$$

where $\Delta\ell_i = \ell_i - \ell_{i-1}$. The discrete maximum entropy choice for $\boldsymbol{q}_i$ solves

$$\max_{\boldsymbol{q}_i} E = -\sum_{j=1}^{N} q_{ij} \log(q_{ij}) \tag{17}$$

subject to

$$q_{ij} \geqslant 0 \quad \forall j,$$

$$\sum_{j=1}^{N} q_{ij} = \pi_i,$$

and

$$\sum_{j=1}^{N} q_{ij} m_{ij} = \pi_i \mu_i.$$

In the ith income range, the discrete maximum entropy probability distribution is

$$q_{ij} = \frac{\pi_i \mathrm{e}^{-\eta_i m_{ij}}}{\sum_{j=1}^{N} \mathrm{e}^{-\eta_i m_{ij}}}, \quad j = 1, \ldots, N, \tag{18}$$

with the Lagrange multiplier for the ith conditional mean constraint, η_i, chosen to satisfy

$$\frac{\sum_{j=1}^{N} m_{ij}\, \mathrm{e}^{-\eta_i m_{ij}}}{\sum_{k=1}^{N} \mathrm{e}^{-\eta_i m_{ik}}} = \pi_i \mu_i. \tag{19}$$

[7] Maximum entropy is motivated, derived, and discussed in detail in Csiszár (1991), Gokhale and Kullback (1978), Golan (1994), Golan et al. (1996a), Jaynes (1957a, b, 1984), Kullback (1959), Kullback and Leibler (1951), and Shannon (1948). Golan et al. (2001), Tobias and Zellner (2001); Zellner (1997a, b); and Zellner et al. (1997) present further developments and applications of maximum entropy to a range of problems. Previous work most like that developed here can be found in Mead and Papanicolaou (1984) and Ryu (1993). An alternative method based on entropy and summary statistics for inferring the distribution of firm size in an industry appears in Golan et al. (1996b).

Defining $f_{ij}(m) = q_{ij}/(\Delta\ell_i/N)\ \forall i = 1,\ldots,n,\ \forall j = 1,\ldots,N$, and $\forall m \in [\ell_{i-1}, \ell_i)$, this generates a piecewise uniform maximum entropy density function for the U.S. income distribution,

$$f(m) = \frac{\pi_i e^{-\eta_i m_{ij}}}{(\Delta\ell_i/N)\sum_{j=1}^{N} e^{-\eta_i m_{ij}}} \quad \forall m \in [\ell_{i-1} + (j-1)\ell_i/N,\ \ell_{i-1} + j\ell_i/N),$$
$$\forall j = 1,\ldots,N,\ \forall i = 1,\ldots,n. \tag{20}$$

Implementation of maximum entropy in specific applications often raises some questions. In our application to the U.S. income distribution, there are two questions of particular interest. The first arises from the fact that each choice of the number of subintervals within each income range generates a different solution for the income distribution. For 1997, Fig. 1a depicts the extreme case of the series of piecewise uniform density functions (the series of horizontal segments in the figure) with $N = 2$ on each income range,

$$f_i(m) = \frac{\pi_i}{\Delta\ell_i/2} \times \begin{cases} \dfrac{(\ell_{i-1} + \Delta\ell_i/4) - \mu_i}{\Delta\ell_i/2}, & m \in [\ell_{i-1}, (\ell_{i-1} + \ell_i)/2), \\ \dfrac{\mu_i - (\ell_{i-1} + 3\Delta\ell_i/4)}{\Delta\ell_i/2}, & m \in [(\ell_{i-1} + \ell_i)/2, \ell_i), \end{cases}$$
$$i = 1,\ldots,6, \tag{21}$$

where we define $\ell_0 = 0$ and we choose $\ell_6 = \ell_5 + 2(\mu_6 - \ell_5)$, which maximizes entropy with respect to ℓ_6.

The opposite extreme for the number of subintervals in the discrete case is to let $N \to \infty$. This gives a continuous conditional density function within each income range. To derive this limiting distribution, for any $s \in [0,1]$ let $[sN]$ denote the integer part of sN, for each $j = 1,\ldots,N$, define the step function $p_i(s) \equiv p_{i[sN]}$, and for $1 \leqslant j \leqslant N$, let s satisfy $(j-1)/N \leqslant s \leqslant j/N$. Then, uniformly in $s \in [0,1]$, the ith conditional cumulative probability distribution function satisfies

$$\begin{aligned} F_{iN}(s) &\equiv \frac{\sum_{k=1}^{j} (\frac{1}{N}) e^{-\eta_i m_{ik}}}{\sum_{k=1}^{N} (\frac{1}{N}) e^{-\eta_i m_{ik}}} \\ &\equiv \frac{\sum_{k=1}^{[sN]} (\frac{1}{N}) \exp\{-\eta_i(\frac{k}{N}\Delta\ell_i + \ell_{i-1} - \frac{1}{2N}\Delta\ell_i)\}}{\sum_{k=1}^{N} (\frac{1}{N}) \exp\{-\eta_i(\frac{k}{N}\Delta\ell_i + \ell_{i-1} - \frac{1}{2N}\Delta\ell_i)\}} \\ &\equiv \frac{\int_0^{[sN]/N} \exp\{-\eta_i[\frac{[xN]}{N}\Delta\ell_i + \ell_{i-1} - (\frac{1}{2N})\Delta\ell_i]\}\,dx}{\int_0^1 \exp\{-\eta_i[\frac{[xN]}{N}\Delta\ell_i + \ell_{i-1} - (\frac{1}{2N})\Delta\ell_i]\}\,dx} \\ &\xrightarrow[N\to\infty]{} \frac{\int_0^s \exp\{-\eta_i[x\Delta\ell_i + \ell_{i-1}]\}\,dx}{\int_0^1 \exp\{-\eta_i[x\Delta\ell_i + \ell_{i-1}]\}\,dx}. \end{aligned} \tag{22}$$

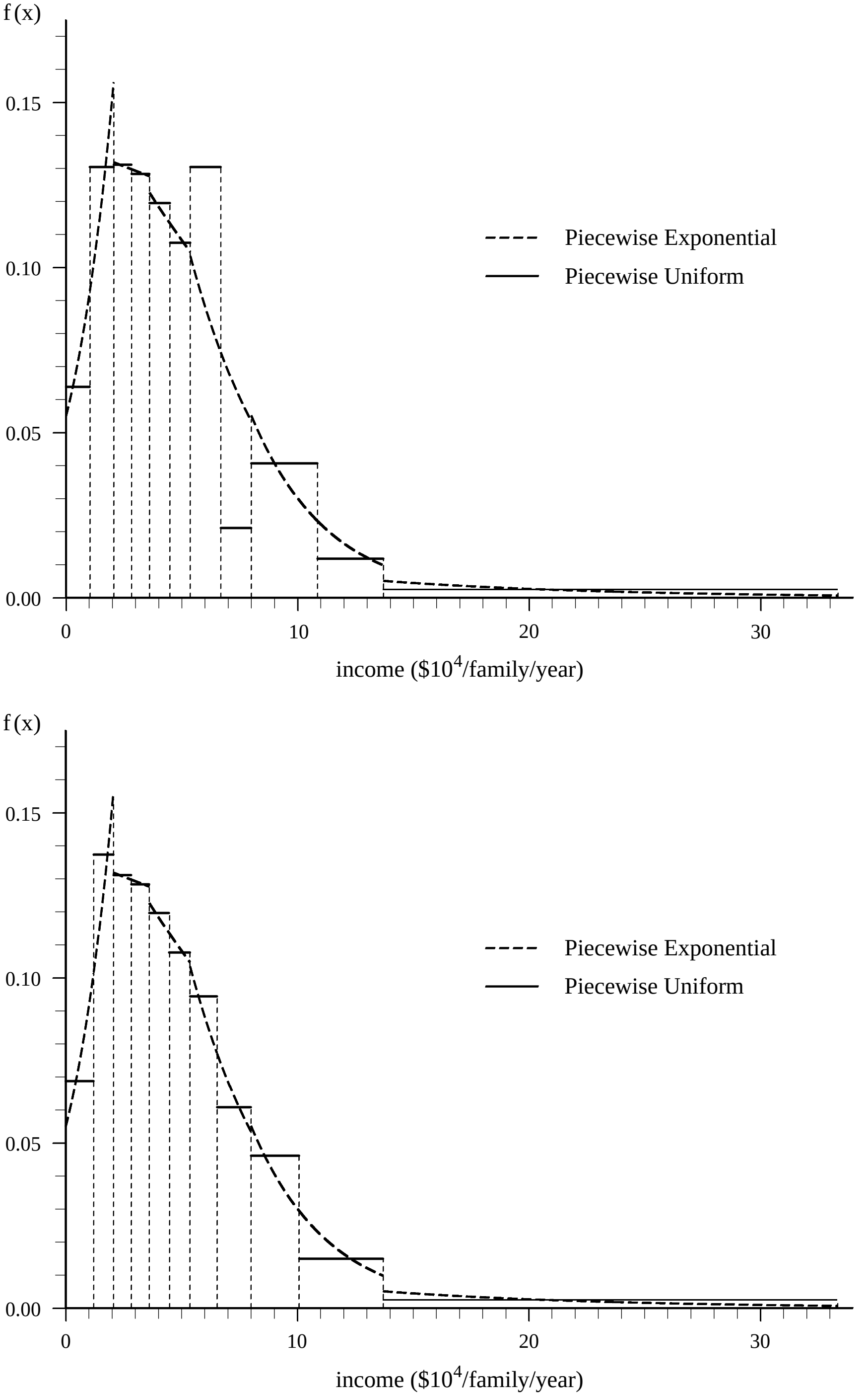
f(x)
0.15
0.10
0.05
0.00
0
10
20
30
Piecewise Exponential
Piecewise Uniform
income ($10^4/family/year)
f(x)
0.15
0.10
0.05
0.00
0
10
20
30
Piecewise Exponential
Piecewise Uniform
income ($10^4/family/year)

Fig. 1.

Changing variables from $s \in [0,1)$ to $m = \ell_{i-1} + \Delta\ell_i s \in [\ell_{i-1}, \ell_i)$ gives the limiting conditional income distribution for the ith income range, given $m \in [\ell_{i-1}, \ell_i)$, as

$$F_i(m) = \frac{\int_{\ell_{i-1}}^{m} \mathrm{e}^{-\eta_i x}\,\mathrm{d}x}{\int_{\ell_{i-1}}^{\ell_i} \mathrm{e}^{-\eta_i x}\,\mathrm{d}x} = \frac{\mathrm{e}^{-\eta_i m} - \mathrm{e}^{-\eta_i \ell_{i-1}}}{\mathrm{e}^{-\eta_i \ell_i} - \mathrm{e}^{-\eta_i \ell_{i-1}}} \quad \forall m \in [\ell_{i-1}, \ell_i), \tag{23}$$

which is a truncated exponential cumulative distribution function. The associated unconditional probability density function is

$$f_i(m) = \frac{-\pi_i \eta_i \mathrm{e}^{-\eta_i m}}{e^{-\eta_i \ell_i} - e^{-\eta_i \ell_{i-1}}} \quad \forall m \in [\ell_{i-1}, \ell_i), \tag{24}$$

and the Lagrange multipliers now satisfy the mean conditions

$$\frac{\int_{\ell_{i-1}}^{\ell_i} -\eta_i m \mathrm{e}^{-\eta_i m}\,\mathrm{d}m}{\mathrm{e}^{-\eta_i \ell_i} - \mathrm{e}^{-\eta_i \ell_{i-1}}} = \left(\frac{\ell_i \mathrm{e}^{-\eta_i \ell_i} - \ell_{i-1} \mathrm{e}^{-\eta_i \ell_{i-1}}}{\mathrm{e}^{-\eta_i \ell_i} - \mathrm{e}^{-\eta_i \ell_{i-1}}} \right) - \left(\frac{1}{\eta_i} \right) = \mu_i. \tag{25}$$

In the empirical application, we exploit the above limiting argument to incorporate closed form numerical expressions for the cross-sectional means of the PIGL functions of income $\mathrm{E}(m^{1-\kappa})$ and $\mathrm{E}(m^{1+\kappa})$ within the demand system estimation procedure.

We next show that this is indeed the continuous maximum entropy solution for the income distribution (also, see Golan et al. (1996, p. 40), Mead and Papanicolaou (1984), and Ryu (1993)). The starting point in the continuous case is the maximum entropy criterion,

$$\max_{\{f\}} E = -\int_0^\infty f(x) \ln(f(x))\,\mathrm{d}x, \tag{26}$$

subject to the intra-range probability and mean constraints for each quintile and the top five percentile of family incomes,

$$\int_{\ell_{i-1}}^{\ell_i} f(x)\,\mathrm{d}x = \pi_i, \quad i = 1, \ldots, 6, \tag{27}$$

$$\int_{\ell_{i-1}}^{\ell_i} x f(x)\,\mathrm{d}x = \pi_i \mu_i, \quad i = 1, \ldots, 6. \tag{28}$$

The continuous maximum entropy solution is the unique proper density function on $\mathbb{R}_+$ that minimizes the average logarithmic height of the density while satisfying the integration and moment constraints. In this respect, it is the density function that is closest to a uniform density, where close is defined by the Kullback–Leibler cross entropy pseudo-distance (Golan et al., 1996; Kullback and Leibler, 1951).

To derive the continuous maximum entropy density function, we seek the maximum with respect to $\{f\}$ and minimum with respect to the Lagrange multipliers $\{\eta, \omega\}$ for the Lagrangean function,

$$\begin{aligned}\mathscr{L} &= -\int_0^{\infty} f(x)\ln(f(x))\,\mathrm{d}x + \sum_{i=1}^{6}\omega_i\left[\pi_i - \int_{\ell_{i-1}}^{\ell_i} f(x)\,\mathrm{d}x\right] \\ &\quad + \sum_{i=1}^{6}\eta_i\left[\pi_i\mu_i - \int_{\ell_{i-1}}^{\ell_i} xf(x)\,\mathrm{d}x\right] \\ &= -\sum_{i=1}^{6}\int_{\ell_{i-1}}^{\ell_i} f(x)[\ln(f(x)) + \eta_i + \omega_i x]\,\mathrm{d}x + \sum_{i=1}^{6}\pi_i(\eta_i + \omega_i\mu_i). \end{aligned} \tag{29}$$

This Lagrangean involves a sum of integrals in $\{f\}$ with no differential or boundary value constraints. It follows from optimal control theory that the necessary (and in this case sufficient) condition for a maximum of $\mathscr{L}$ with respect to $\{f\}$ is the point-wise maximum of each of the six individual integrands with respect to f. Upon carrying out the necessary steps, we obtain a sequence of six exponential densities, each defined on their respective ranges,

$$f(m) = \begin{cases} \dfrac{-\pi_i\eta_i \mathrm{e}^{-\eta_i m}}{\mathrm{e}^{-\eta_i\ell_i} - \mathrm{e}^{-\eta_i\ell_{i-1}}}, & m\in[\ell_{i-1},\ell_i),\ \ i=1,\ldots,5, \\ \dfrac{\pi_6\mathrm{e}^{-(m-\ell_5)/(\mu_6-\ell_5)}}{\mu_6-\ell_5}, & m\in[\ell_5,\infty), \end{cases} \tag{30}$$

with $\pi_i = 0.20$, for $i = 1,\ldots,4$, $\pi_5 = 0.15$ and $\pi_6 = 0.05$. This is indeed the same sequence of piecewise exponential densities in (24) with the Lagrange multipliers for the mean constraints satisfying (25), which can be rewritten for the purpose of numerical solution as,[8]

$$e^{\eta_i(\ell_i-\ell_{i-1})} - \left[\frac{1+\eta_i(\ell_i-\mu_i)}{1-\eta_i(\mu_i-\ell_{i-1})}\right] = 0, \quad i=1,\ldots,5. \tag{25$'$}$$

For 1997, this solution for the income distribution is depicted both in Fig. 1a and b as the sequence of six piecewise exponential segments, continuous within each income range, and discontinuous at the intra-range boundaries.

Returning to the discrete case briefly, the second question regarding the implementation of the maximum entropy methodology is where to locate the discontinuity points within each income range. We illustrate the impact of this by constructing a piecewise uniform conditional density function on each range that is discontinuous at the

[8] The unique solution for each η_i is readily obtained from the following observations. If $\mu_i = (\ell_i + \ell_{i-1})/2$ there is only one solution, $\eta_i = 0$. Second, the first term in (25′) is always positive, while the second term has a zero at $\eta_i = -(\ell_i - \mu_i)^{-1} < 0$ and a pole at $\eta_i = (\mu_i - \ell_{i-1})^{-1} > 0$. If $\mu_i > (\ell_i + \ell_{i-1})/2$, then $\eta_i < 0$ and interval halving over $\eta_i \in (-(\ell_i - \mu_i)^{-1}, 0)$ produces a numerical approximation with accuracy on the order of 2^{-k}, where k is the number of interval squeezes. If $\mu_i < (\ell_i + \ell_{i-1})/2$, $\eta_i > 0$ and successive interval halving on $\eta_i \in (0, (\mu_i - \ell_{i-1})^{-1})$ produces the desired approximation. In the empirical application, the numerical estimates always converged (in double precision, to fifteen digits) after less than twenty iterations.

intra-range conditional mean, μ_i, rather than at the midpoint of the corresponding income range. This generates a piecewise uniform density function defined by

$$f_i(m) = \frac{\pi_i}{\Delta \ell_i} \times \begin{cases} \dfrac{\ell_i - \mu_i}{\mu_i - \ell_{i-1}} \ \forall m \in (\ell_{i-1}, \mu_i], \\ \dfrac{\mu_i - \ell_{i-1}}{\ell_i - \mu_i} \ \forall m \in (\mu_i, \ell_i], \end{cases} \qquad i = 1, \ldots, 6, \tag{31}$$

again with $\ell_0 = 0$ and $\ell_6 = \ell_5 + 2(\mu_6 - \ell_5)$. Fig. 1b depicts this solution for the 1997 U.S. income distribution as the series of horizontal segments.

It is interesting to note that, even with $N = 2$, setting the discontinuity points at the conditional mean leads to a distribution whose overall shape is considerably closer to the piecewise exponential relative to the piecewise uniform distribution with discontinuities at the midpoints. Of course, increasing the number of sub-intervals would produce the same effect, ultimately leading to the piecewise exponential density function. Our primary purpose is to use the information theoretic measures of the U.S. income distribution to construct the cross-sectional means $\mathrm{E}(m^{1-\kappa})$ and $\mathrm{E}(m^{1+\kappa})$. We want to assess the robustness of the demand model's parameter estimates to the way in which these income moments are calculated, particularly with respect to the rank and functional form of the income terms in the demand equations. To accomplish this, we might contrast the results obtained from the two income distributions that are least similar. Therefore, in our empirical application, we compare the piecewise uniform with $N = 2$ and discontinuities at the intra-range midpoints with the piecewise exponential.

3.1. Extrapolating the income distribution 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, 1935–1936, 1941 and 1946–1998. One issue that arises in using this data in an aggregate U.S. food demand model, then, centers on predicting the missing income data for the years 1918–1928, 1930–1940, 1942–1943, and 1945.

To forecast the upper limit of the first quintile, $\ell_{1,t}$, we use per capita disposable personal income and the unemployment rate as predictors. We estimate a least squares relationship with the log of the first quintile's upper 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 autocorrelation in the error terms,

$$\begin{aligned} \ln(\ell_{1,t}) &= \alpha + \beta_1 \ln(\mu_{\mathrm{pc},t}) + \beta_2[\ln(\mu_{\mathrm{pc},t})]^2 + \beta_3 u_t + e_{1,t}, \\ e_{1,t} &= \rho e_{1,t-1} + \varepsilon_{1,t}, \end{aligned} \tag{32}$$

where $\mu_{\mathrm{pc},t}$ is per capita disposable personal income and u_t is the annual average U.S. unemployment rate. Missing values for $\ln(\ell_{1,t})$ are replaced with predictions from this regression equation for the years 1918–1928, 1930–1940, 1942–1943, and 1945. 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,

$$\ln(\ell_{i,t}) = \alpha_i + \beta_{i1}\ln(\ell_{i-1,t}) + \beta_{i2}[\ln(\ell_{i-1,t})]^2 + \beta_{i3}[\ln(\ell_{i-1,t})]^3 + e_{i,t}, \tag{33}$$

for $i = 2, \ldots, 5$. We replace the missing observations for the income limit variables in each case with the least squares forecast obtained from these regression equations.

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,

$$\ln(\mu_{1,t}) = \alpha + \beta_1\ln(\mu_{\mathrm{pc},t}) + \beta_2[\ln(\mu_{\mathrm{pc},t})]^2 + \beta_3 u_t + v_{1,t},$$
$$v_{1,t} = \rho v_{1,t-1} + \upsilon_{1,t}. \tag{34}$$

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,

$$\ln(\mu_{i,t}) = \alpha_i + \beta_{i1}\ln(\mu_{i-1,t}) + \beta_{i2}[\ln(\mu_{i-1,t})]^2 + \beta_{i3}[\ln(\mu_{i-1,t})]^3 + \beta_{i4}\ln(\mu_{\mathrm{pc},t}) + \beta_{i5}u_t + v_{i,t}, \tag{35}$$

for $i=2,\ldots,6$. As before, we replace the missing observations for the intra-range means with the least squares predictions. The summary statistics presented Table 1 and visual inspection of plots of the observed data and regression curves (see LBPA) suggest that these conditional prediction equations are very precise, indicating perhaps a system of nonlinear cointegrating relationships.

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 demand for the years 1918–1995, excluding 1942–1946, can be written in deflated expenditure form as

$$\begin{aligned} \boldsymbol{e}_t = m_t^{1-\kappa}\boldsymbol{P}_t^{\lambda}\{&\boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{B}\boldsymbol{p}_t(\lambda) + \boldsymbol{\gamma}[m_t(\kappa) - \boldsymbol{p}(\lambda)'\boldsymbol{A}\boldsymbol{s}_t - \tfrac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda)] \\ &+ [\boldsymbol{I} + \boldsymbol{\gamma}\boldsymbol{p}_t(\lambda)']\boldsymbol{\delta}[m_t(\kappa) - \boldsymbol{p}_t(\lambda)'\boldsymbol{A}\boldsymbol{s}_t - \tfrac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda)]^2\} + \boldsymbol{\varepsilon}_t, \\ &t = 1, \ldots, T, \end{aligned} \tag{36}$$

Table 1
Prediction equations for the U.S. income distribution

Upper income range limits							
Var.	Const.	$\ln(\mu_{pc})$	$[\ln(\mu_{pc})]^2$	Unemp.	ρ	R^2	D–W
$\ln(\ell_1)$	−5.568	2.514	−0.09593	−1.439	0.6529	0.9990	1.621
	(0.9566)	(0.2278)	(0.01332)	(0.3638)	(0.1093)		
Var.	Const.	$\ln(\ell_{i-1})$	$[\ln(\ell_{i-1})]^2$	$[\ln(\ell_{i-1})]^3$		R^2	D–W
$\ln(\ell_2)$	−2.088	2.101	−0.1516	0.006813		0.9994	1.459
	(1.942)	(0.7209)	(0.08854)	(0.003597)			
$\ln(\ell_3)$	7.242	1.177	0.2223	−0.007330		0.9999	1.237
	(1.178)	(0.4093)	(0.0471)	(0.001794)			
$\ln(\ell_4)$	5.674	−0.4928	0.1331	−0.003720		0.9999	1.067
	(1.151)	(0.3824)	(0.04210)	(0.001535)			
$\ln(\ell_5)$	13.21	−2.569	0.3267	−0.009746		0.9993	1.523
	(3.505)	(1.113)	(0.1173)	(0.004095)			

Conditional income means								
Var.	Const.	$\ln(\mu_{pc})$		$[\ln(\mu_{pc})]^2$	Unemp.	ρ	R^2	D–W
$\ln(\mu_1)$	−5.719	2.429		−0.09083	−1.982	0.8943	0.9978	1.621
	(2.819)	(0.6691)		(0.03915)	(0.5293)	(0.0593)		
	Const.	$\ln(\mu_{i-1})$	$[\ln(\mu_{i-1})]^2$	$[\ln(\mu_{i-1})]^3$	$\ln(\mu_{pc})$	Unemp.	R^2	D–W
$\ln(\mu_2)$	7.286	3.190	−0.3061	0.01057	0.5804		0.9998	1.251
	(1.275)	(0.5510)	(0.0721)	(0.00325)	(0.0336)			
$\ln(\mu_3)$	0.9981	0.7409	0.00888		0.1165	0.3149	0.9998	1.351
	(0.4022)	(0.0577)	(0.00623)		(0.0542)	(0.0855)		
$\ln(\mu_4)$	2.661	0.5136	0.03200		−0.07413		0.9999	1.056
	(0.2268)	(0.0299)	(0.00352)		(0.03695)			
$\ln(\mu_5)$	27.92	−6.172	0.6920	−0.01980	0.5699	−1.399	0.9990	1.365
	(5.472)	(1.804)	(0.1899)	(0.00679)	(0.1390)	(0.2621)		
$\ln(\mu_6)$	6.084	0.2495	0.06530		−0.5704	−0.5396	0.9997	1.444
	(0.3858)	(0.0878)	(0.00333)		(0.0323)	(0.1020)		

Numbers in parentheses are estimated standard errors.
D–W is the Durbin–Watson statistic for serial correlation.

where $\boldsymbol{e}_t = [p_{1t}q_{1t} \dots p_{nt}q_{nt}]'$ is the 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 $\boldsymbol{\varepsilon}_t$ is a 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 Eq. (36) and grouping terms, the QPIGL–IDS demand equations can be rewritten in the form

$$\begin{aligned}\boldsymbol{e}_t = \boldsymbol{P}_t^{\lambda}\Bigg\{&\Bigg[\boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{B}\boldsymbol{p}_t(\lambda) - \gamma\left(\frac{1}{\kappa} + \boldsymbol{p}(\lambda)'\boldsymbol{A}\boldsymbol{s}_t + \frac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda)\right)\\ &+ [\boldsymbol{I} + \gamma\boldsymbol{p}_t(\lambda)']\delta\left(\boldsymbol{p}_t(\lambda)'\boldsymbol{A}\boldsymbol{s}_t + \frac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda) + \frac{1}{\kappa}\right)^2\Bigg] m_t^{1-\kappa}\\ &+ \frac{1}{\kappa}\left[\gamma - 2[\boldsymbol{I} + \gamma\boldsymbol{p}_t(\lambda)']\delta\left(\boldsymbol{p}_t(\lambda)'\boldsymbol{A}\boldsymbol{s}_t + \frac{1}{2}\boldsymbol{p}_t(\lambda)'\boldsymbol{B}\boldsymbol{p}_t(\lambda) + \frac{1}{\kappa}\right)\right] m_t\\ &+ \left[\frac{\boldsymbol{I} + \gamma\boldsymbol{p}_t(\lambda)'}{\kappa^2}\right]\delta m_t^{1+\kappa}\Bigg\} + \boldsymbol{\varepsilon}_t, \qquad (37)\end{aligned}$$

which shows explicitly how the three income variables $m_t^{1-\kappa}$, m_t, and $m_t^{1+\kappa}$ enter the demand model. Consistently estimating the model's parameters using aggregate market data therefore requires, for any value of κ, evaluation of the cross-sectional means 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.

We use two-step nonlinear seemingly unrelated regressions equations (NLSURE) estimation methods. 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. [9] 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. Initially, during each iteration on the residual variance–covariance matrix, a one-dimensional search over κ is used to ensure that we obtain the global minimum of the residual sum of squares. We follow this with joint estimation of all parameters, incorporating

[9] See LaFrance (1999b) for a discussion. The crux of the matter is that all 317 model parameters enter each of the demand equations, with only 76 time series observations. This has the possibility of a numerically singular estimated residual covariance matrix when iterative NLSURE is employed.

numerical approximations to the income moment conditions within the final model estimates.[10]

Fig. 2 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 matrix fixed at the value calculated from the total sum of squared residuals minimizing value for κ. The first round values for κ are 0.983 and 0.952 for the piecewise exponential and uniform income distributions, respectively, while the second round values are 0.983 and 0.955, respectively. As can be seen in the figure (and in the first line of Table 3), the QAIDS–IDS form of the income terms is rejected in favor of a QPIGL–IDS that is numerically very close to an extended QES–IDS. This result holds for either form for the income distribution and for either traditional (white noise) or robust (heteroskedasticity consistent) estimated standard errors, at any reasonable level of statistical significance. The estimate for the Box–Cox price parameter is 0.739 and 0.794 for the piecewise exponential and uniform distributions, respectively. These point estimates, along with their corresponding estimated standard errors, reject both the logarithmic and the linear forms for prices at any reasonable level of statistical significance.

The results of Wald tests for the rank of the demand model, with p-values in square brackets below the chi-square point estimates, are:

Null	Degrees of freedom	Exponential		Uniform	
		Traditional	Robust	Traditional	Robust
$\gamma = 0$	21	108.1	245.4	5207	25,775
		[0.000]	[0.000]	[0.000]	[0.000]
$\boldsymbol{\delta} = 0$	21	30.63	79.07	70.01	202.78
		[0.080]	[0.000]	[0.000]	[0.000]
$\gamma = \boldsymbol{\delta} = 0$	42	332.3	1310	6737	56,558
		[0.000]	[0.000]	[0.000]	[0.000]

[10] The piecewise uniform distribution admits simple closed form solutions for these expectations of the generic form $E(m^{1\pm\kappa}) = \sum_{i,j} f_{ij}(x_{ij}^{2\pm\kappa} - x_{ij-1}^{2\pm\kappa})/(2 \pm \kappa)$. For the piecewise exponential, in the lower four income quintiles and the 80–95 percentile, the generic form of the density function is $f(x) = \alpha e^{\beta x}, \forall x \in [\ell, \upsilon)$, with $\alpha = \pi\beta/(e^{\beta\upsilon} - e^{\beta\ell})$, where π is the proportion of families with income in the given range. We break each range into N equally spaced subintervals with limits $\ell_i = \ell + i(\upsilon - \ell)/N$, $i = 1, \ldots, N$, $\ell_0 = \ell$ and $\ell_N = \upsilon$. The probability that x is in the ith sub-interval is $q_i = \alpha(e^{\beta\ell_i} - e^{\beta\ell_{i-1}})/\beta$, $i = 1, \ldots, N$. We use the q_i as probability weights and evaluate x at each sub-interval midpoint, $x_i = (\ell_{i-1} + \ell_i)/2$. Because the top five percentile is unbounded, the probability weights and evaluation points are different. The density function in this range has the generic form $f(x) = \pi e^{-(x-\ell)/(\mu-\ell)}/(\mu - \ell)\ \forall x \in [\ell, \infty)$, where ℓ is the lower limit, μ is the conditional mean, and π is the proportion of families with incomes in the top 5%. We use equal probabilities (i.e., $q_i = \pi/N$) to define sub-intervals with upper limits, $\ell_i = \ell + (\mu - \ell)\ln(N/(N-i))$, for $i = 1, \ldots, N-1$, and evaluate x at the sub-interval midpoints. For the last subinterval, $[\ell_{n-1}, \infty)$, we evaluate x at the conditional mean for this sub-interval, $x_N = \mu_N = \ell_N + \mu - \ell = \mu + (\mu - \ell)\ln(N)$. In all of these cases, we have $\sum_{i=1}^{N} q_i x_i^{1\pm\kappa} \underset{N\to\infty}{\longrightarrow} \int_{\ell}^{\upsilon} x^{1\pm\kappa} f(x)\,dx$. In our empirical application, $n = 30$ gives excellent results.

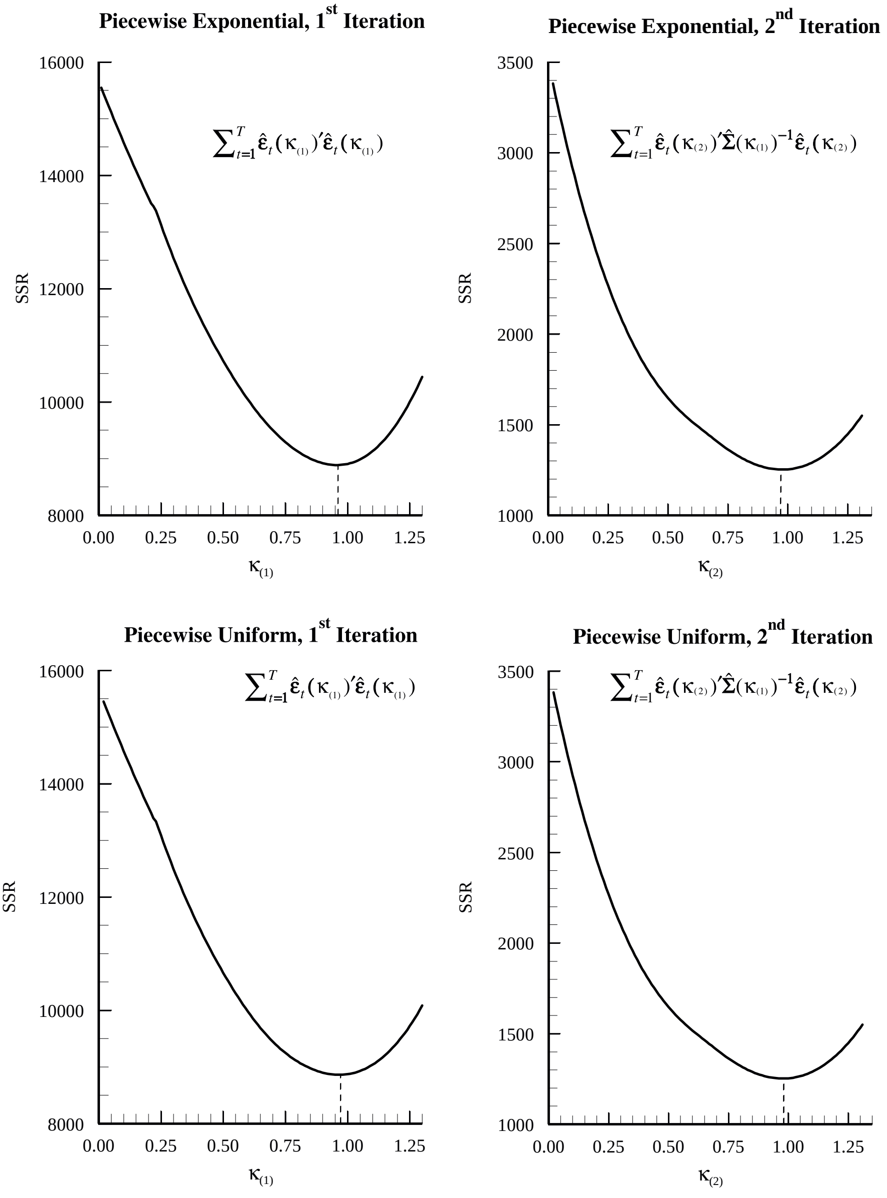

Fig. 2.

We reject all three null hypotheses at the 10 percent level of significance. We also reject the hypothesis that the linear income coefficients vanish, as well as the joint hypothesis that all linear and quadratic income coefficients all vanish, at any reasonable level of significance. However, we do not reject the hypothesis that the quadratic

Table 2
Equation summary statistics

	Dependent variable		Piecewise exponential		Piecewise uniform	
Food group	Mean	S.D.	R^2	D–W	R^2	D–W
Fresh milk & cream	128.9	34.76	0.9975	1.696	0.9968	1.759
Butter	30.69	22.31	0.9949	1.055	0.9938	1.092
Cheese	41.03	23.44	0.9979	1.705	0.9981	1.451
Frozen dairy	15.01	4.046	0.9686	1.567	0.9678	1.609
Canned & dried milk	11.84	3.618	0.9721	1.995	0.9707	2.067
Beef & veal	240.8	80.09	0.9917	1.335	0.9923	1.400
Pork	126.5	27.02	0.9763	1.205	0.9739	1.101
Other red meat	36.17	8.715	0.9616	1.554	0.9644	1.570
Fish	28.52	10.50	0.9911	1.616	0.9891	1.680
Poultry	56.13	12.45	0.9819	1.128	0.9834	1.185
Fresh citrus	16.42	2.729	0.7770	1.597	0.7612	1.522
Fresh noncitrus	43.36	19.37	0.9585	2.746	0.9615	2.768
Fresh vegetables	58.18	10.05	0.9768	1.955	0.9731	1.940
Potatoes	24.46	6.793	0.9771	2.581	0.9769	2.564
Processed fruit	83.23	37.84	0.9864	1.446	0.9872	1.462
Processed vegetables	40.12	8.048	0.9821	1.435	0.9787	1.293
Eggs	48.54	7.562	0.9809	1.922	0.9809	1.738
Fats & oils	58.29	29.95	0.9982	1.741	0.9980	1.756
Cereals & bakery	71.89	15.03	0.9888	1.398	0.9894	1.586
Sweeteners	93.07	19.35	0.9861	2.129	0.9858	2.222
Coffee, tea & cocoa	45.06	12.37	0.9779	1.630	0.9777	1.566

income coefficients vanish for the case of the piecewise exponential income distribution and traditional estimated standard errors at the 5 percent level. Even so, we find the sum of the evidence to be compelling against a rank two demand model. We also find uniformly compelling evidence against any form of AIDS–IDS. For the piecewise exponential income distribution, we do not reject the hypothesis that the QPIGL–IDS has a Box–Cox income parameter of one at standard levels of significance. However, we reject this hypothesis at the 1 percent level for the piecewise uniform income distribution, although the point estimate for the Box–Cox income parameter is numerically close to one.

Table 2 presents equation summary statistics for both versions of the income distribution and Table 3 presents the Box–Cox price (λ) and income (κ) coefficients and the linear (γ_1–γ_{21}) and quadratic (δ_1–δ_{21}) income coefficients. In these tables, the Box–Cox parameters are estimated jointly with the other model parameters and the standard errors are calculated using both traditional (white noise) and robust (heteroskedasticity consistent) formulas. Statistical fits and general lack of evidence for substantial serial correlation in the error terms, suggests at first blush that the two procedures for constructing the income distribution perform similarly. However, the point estimates and standard errors for the linear and quadratic income coefficients tell a different story. The magnitude and signs of the income coefficients are very different in the two specifications for the income distribution. The signs of the income coefficients are reversed

Table 3
Functional form and rank income parameter estimates

	Piecewise Exponential			Piecewise Uniform		
		Standard Error			Standard Error	
Param.	Estimate	Gauss	Robust	Estimate	Gauss	Robust
κ	0.983	0.0186*	0.0158*	0.952	0.0167*	0.0141*
λ	0.739	0.0285*	0.0303*	0.794	0.0245*	0.0257*
γ_1	0.0278	0.0109*	0.0103*	−0.789	0.0696*	0.0647*
γ_2	0.00146	0.00253	0.00180	−0.100	0.0127*	0.0088*
γ_3	0.00439	0.00191*	0.00140*	0.0182	0.0131	0.0136
γ_4	0.00699	0.00459	0.00269*	−0.122	0.0448*	0.0353*
γ_5	−0.00231	0.00417	0.00238	−0.135	0.0385*	0.0324*
γ_6	0.0171	0.00640*	0.00740*	0.0893	0.0133*	0.0133*
γ_7	0.0224	0.00527*	0.00569*	0.0558	0.0167*	0.0147*
γ_8	0.00254	0.00281	0.00204	−0.0464	0.0186*	0.0210*
γ_9	0.00252	0.00170	0.00137	−0.0507	0.0123*	0.0117*
γ_{10}	0.00138	0.00328	0.00343	−0.0996	0.0165*	0.0168*
γ_{11}	0.00932	0.00519	0.00483	−0.110	0.0252*	0.0207*
γ_{12}	0.0239	0.0139	0.0122*	0.304	0.0468*	0.0493*
γ_{13}	0.00631	0.00597	0.00616	−0.174	0.0427*	0.0338*
γ_{14}	−0.00287	0.00882	0.00738	−0.139	0.0507*	0.0518*
γ_{15}	0.00537	0.00633	0.00693	0.178	0.0207*	0.0162*
γ_{16}	0.0225	0.00518*	0.00307*	−0.059	0.0422	0.0354
γ_{17}	0.0150	0.00291*	0.00232*	0.137	0.0212*	0.0184*
γ_{18}	0.0116	0.00324*	0.00292*	0.174	0.0184*	0.0209*
γ_{19}	−0.0135	0.00964	0.00916	−0.348	0.0707*	0.0858*
γ_{20}	0.0156	0.00681*	0.00688*	−0.0816	0.0205*	0.0178*
γ_{21}	0.00514	0.00202*	0.00154*	0.0177	0.00538*	0.00381*
δ_1	-7.67×10^{-7}	4.09×10^{-7}	3.39×10^{-7}*	1.60×10^{-6}	3.46×10^{-7}*	3.68×10^{-7}*
δ_2	-3.36×10^{-8}	1.05×10^{-7}	7.43×10^{-8}	1.98×10^{-7}	5.21×10^{-8}*	4.03×10^{-8}*
δ_3	-1.61×10^{-7}	8.0×10^{-8}*	5.77×10^{-8}*	-3.60×10^{-8}	2.68×10^{-8}	2.64×10^{-8}
δ_4	-2.89×10^{-7}	1.98×10^{-7}	1.15×10^{-7}*	2.40×10^{-7}	9.49×10^{-8}*	9.36×10^{-8}*
δ_5	1.74×10^{-7}	1.68×10^{-7}	1.01×10^{-7}	2.80×10^{-7}	9.45×10^{-8}*	6.84×10^{-8}*
δ_6	-1.93×10^{-7}	2.52×10^{-7}	2.77×10^{-7}	-1.11×10^{-7}	3.96×10^{-8}*	3.89×10^{-8}*
δ_7	-5.20×10^{-7}	2.33×10^{-7}*	2.19×10^{-7}*	-5.61×10^{-8}	3.41×10^{-8}	3.34×10^{-8}
δ_8	-7.16×10^{-9}	1.12×10^{-7}	8.69×10^{-8}	9.37×10^{-8}	4.16×10^{-8}*	4.66×10^{-8}*
δ_9	-2.18×10^{-8}	6.52×10^{-8}	5.32×10^{-8}	1.11×10^{-7}	3.54×10^{-8}*	3.60×10^{-8}*
δ_{10}	2.64×10^{-7}	1.53×10^{-7}	1.70×10^{-7}	2.36×10^{-7}	5.20×10^{-8}*	4.47×10^{-8}*
δ_{11}	-1.85×10^{-7}	2.04×10^{-7}	1.80×10^{-7}	2.43×10^{-7}	7.19×10^{-8}*	6.80×10^{-8}*
δ_{12}	-8.66×10^{-7}	5.92×10^{-7}	4.76×10^{-7}	-5.92×10^{-7}	1.50×10^{-7}*	1.58×10^{-7}*
δ_{13}	-1.48×10^{-7}	2.35×10^{-7}	2.57×10^{-7}	3.53×10^{-7}	1.05×10^{-7}*	8.28×10^{-8}*
δ_{14}	2.25×10^{-7}	3.41×10^{-7}	2.50×10^{-7}	2.76×10^{-7}	1.23×10^{-7}*	1.18×10^{-7}*
δ_{15}	-5.54×10^{-8}	2.65×10^{-7}	3.05×10^{-7}	-3.27×10^{-7}	8.58×10^{-8}*	8.59×10^{-8}*
δ_{16}	-6.04×10^{-7}	2.38×10^{-7}*	1.41×10^{-7}*	1.63×10^{-7}	8.16×10^{-8}*	7.52×10^{-8}*
δ_{17}	-4.71×10^{-7}	1.40×10^{-7}*	9.56×10^{-8}*	-2.47×10^{-7}	6.74×10^{-8}*	5.96×10^{-8}*
δ_{18}	-4.08×10^{-7}	1.54×10^{-7}*	1.16×10^{-7}*	-3.33×10^{-7}	7.93×10^{-8}*	7.67×10^{-8}*
δ_{19}	6.03×10^{-7}	3.83×10^{-7}	3.75×10^{-7}	6.78×10^{-7}	1.81×10^{-7}*	1.43×10^{-7}*
δ_{20}	-3.23×10^{-7}	2.72×10^{-7}	2.44×10^{-7}	1.99×10^{-7}	5.96×10^{-8}*	6.41×10^{-8}*
δ_{21}	-1.05×10^{-7}	8.12×10^{-8}	6.43×10^{-8}	-2.06×10^{-8}	1.20×10^{-8}	8.00×10^{-9}*

*Indicates that the coefficient estimate is statistically significant with this standard error at the 5% level.

almost 50 percent of the time—10 out of 21 possible cases for the linear terms and 9 out of 21 for the quadratic terms. In addition, at the usual 5 percent level of significance, inferences regarding the statistical significance of these income parameters are reversed in 14 and 16 (13 and 14) cases for the linear and quadratic terms, respectively, using traditional (robust) standard errors. [11]

5. Discussion

This paper presented a new method to nest, test and estimate both the rank and functional form of the income terms in an incomplete system of aggregable demand equations. We applied information theory to the problem of inferring the U.S. income distribution from annual time series data on quintile and top five percentile income ranges and intra-quintile and top five percentile mean incomes. These maximum entropy estimates for the year-to-year income distribution were combined with annual time series data on the U.S. consumption of and retail prices for 21 food items over the period 1919–1995, excluding 1942–1946. The empirical results suggest that all versions of the AIDS–IDS are strongly rejected by this data set, in favor of a full rank three QPIGL—IDS that is numerically very close to an extended QES–IDS. This has potentially significant implications for future demand analysis, particularly with respect to food consumption using aggregate market-level data sets. However, the two different versions of the income distribution generated highly conflicting results on the sign, size, and statistical significance of most of the coefficients that interact with income in the demand model.

One interpretation of the empirical results is that the level and form of the prior information that is extracted from the summary statistics on the income distribution significantly affects inferences on food demand. The piecewise exponential is equivalent to a prior belief, or assumption, that the income distribution is continuous. If this is true, or even reasonably accurate, then the piecewise exponential distribution will give a closer approximation to the true income distribution than a piecewise uniform distribution with two segments within each percentile range. On the other hand, if one does not have strong beliefs about whether the income distribution is continuous and desires to introduce absolutely the minimum amount of prior assumptions into the empirical analysis, the piecewise uniform distribution may be preferable.

In either case, it is best to interpret the cross-sectional moments for the QPIGL income terms obtained from the maximum entropy distributions as instruments. Using these instruments increases the power of tests for the rank and functional form of the demand model relative to raising mean income to the powers $1-\kappa$ and $1+\kappa$ and including these in the demand equations. Nevertheless, more information on the form

[11] Estimation of the demand model using the two different income distributions also was quite different. The model with the piecewise exponential income distribution was numerically stable, converging to the minimum sum of squared residuals quickly and with ease once good starting values were obtained using a grid search on κ. In contrast, the model with the piecewise uniform income distribution was far more difficult, requiring substantially more computational time and a finer grid search on κ to obtain estimates for the minimum of the residual sum of squares.

of the income distribution and its changes over time, beyond the quintile and top five percentile ranges and mean incomes, appears to be important to empirical aggregate demand analysis.

References

Anderson, G., Blundell, R., 1983. Testing restrictions in a flexible dynamic demand system: an application to consumer's expenditure in Canada. Review of Economic Studies 50, 397–410.

Buse, A., 1998. Testing homogeneity in the linearized almost ideal demand system. American Journal of Agricultural Economics 80, 208–220.

Csiszár, I., 1991. Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. The Annals of Statistics 19, 2032–2066.

Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. American Economic Review 70, 312–326.

Epstein, L., 1982. Integrability of incomplete systems of demand functions. Review of Economic Studies 49, 411–425.

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

Golan, A., 1994. A multivariable stochastic theory of size distribution of firms with empirical evidence. Advances in Econometrics 10, 1–46.

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

Golan, A., Judge, G., Perloff, J., 1996a. A maximum entropy approach to recovering information from multinomial response data. Journal of the American Statistical Association 91, 841–853.

Golan, A., Perloff, J.M., Shen, Z., 2001. Estimating a demand system with non-negativity constraints: Mexican meat demand. Review of Economics and Statistics, Forthcoming.

Golan, A., Judge, G.G., Perloff, J.M., 1996b. Estimating the size distribution of firms using government summary statistics. Journal of Industrial Economics XLIV, 69–80.

Gorman, W.M., 1995. Consumer budgets and price indices. Unpublished typescript, 1965. Published as Chapter 5. In: Blackorby, C., Shorrocks, T. (Eds.), Separability and Aggregation: Collected Works of W.M. Gorman, Vol. I. Clarendon Press, Oxford, pp. 61–88.

Gorman, W.M., 1981. Some Engel Curves. In: Deaton, A. (Ed.), Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone. Cambridge University Press, Cambridge, 1981. Republished as Chapter 20. In: Blackorby, C., Shorrocks, T. (Eds.), Separability and Aggregation: Collected Works of W.M. Gorman, Vol. I. Clarendon Press, Oxford, pp. 351–376.

Hausman, J., 1981. Exact consumer's surplus and deadweight loss. American Economic Review 71, 662–676.

Howe, H., Pollak, R.A., Wales, T.J., 1979. Theory and time series estimation of the quadratic expenditure system. Econometrica 47, 1231–1247.

Jaynes, E., 1957a. Information theory and statistical mechanics. Physics Review 106, 620–630.

Jaynes, E., 1957b. Information theory and statistical mechanics II. Physics Review 108, 171–190.

Jaynes, E., 1984. Prior information and ambiguity in inverse problems. In: McLaughlin, D. (Ed.), Inverse Problems. American Mathematical Society, Providence RI, pp. 151–166.

Kullback, J., 1959. Information Theory and Statistics. Wiley, New York.

Kullback, J., Leibler, R., 1951. On information and sufficiency. Annals of Mathematical Statistics 4, 99–111.

LaFrance, J.T., 1985. Linear demand functions in theory and practice. Journal of Economic Theory 37 (1), 147–166.

LaFrance, J.T., 1990. Incomplete demand systems and semilogarithmic demand models. Australian Journal of Agricultural Economics 34, 118–131.

LaFrance, J.T., 1999a. An econometric model of the demand for food and nutrition. The 1999 Havlicek Memorial Lecture in Applied Economics, Ohio State University, Columbus OH.

LaFrance, J.T., 1999b. U.S. food and nutrient demand and the effects of agricultural policies. Proceedings of the Food and Agricultural Marketing Consortium, Washington DC.

LaFrance, J., Hanemann, M., 1989. The dual structure of incomplete demand systems. American Journal of Agricultural Economics 71, 262–274.

Lewbell, A., 1987. Characterizing some Gorman Engel curves. Econometrica 55, 1451–1459.
Lewbell, A., 1990. Full rank demand systems. International Economic Review 31, 289–300.
Mead, L.R., Papanicolaou, N., 1984. Maximum entropy in the problem of moments. Journal of Mathematical Physics 25 (8), 2404–2417.
Moschini, G., 1995. Units of measurement and the stone index in demand system estimation. American Journal of Agricultural Economics 77, 63–68.
Moschini, G., Meilke, K., 1989. Modeling the pattern of structural change in U.S. meat demand. American Journal of Agricultural Economics 71, 253–261.
Muellbauer, J., 1975. Aggregation, income distribution and consumer demand. Review of Economic Studies 42, 525–543.
Muellbauer, J., 1976. Community preferences and the representative consumer. Econometrica 44, 979–999.
Pashardes, P., 1993. Bias in estimating the almost ideal demand system with the stone index approximation. Economic Journal 103, 908–915.
Ryu, H.K., 1993. Maximum entropy estimation of density and regression functions. Journal of Econometrics 56 (3), 397–440.
Shannon, C., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423.
Tobias, J., Zellner, A., 2001. Further results on the Bayesian method of moments analysis of the multiple regression model. International Economic Review, Forthcoming.
van Daal, J., Merkies, A.H.Q.M., 1989. A note on the quadratic expenditure model. Econometrica 57, 1439–1443.
Zellner, A., 1997a. The information theoretic maximum entropy: theory and applications. In: Fomby, T., Hill, R. (Eds.), Advances in Econometrics 12, 85–105.
Zellner, A., 1997b. The Bayesian method of moments (BMOM): theory and applications. In: Fomby, T., Hill, R. (Eds.), Advances in Econometrics 12, 85–105.
Zellner, A., Tobias, J., Ryu, K., 1997. Bayesian method of moments (BMOM) analysis of parametric and semiparametric regression models. Alexander Research Foundation, University of Chicago.