School of Economic Sciences

Working Paper Series
WP 2008-10

The Structure of US Food Demand

By

Jeffrey LaFrance

2008

Washington State University
World Class. Face to Face.

The Structure of US Food Demand

Jeffrey T. LaFrance†
Washington State University

In Press: *Journal of Econometrics*, 2008.

Abstract: An exactly aggregable system of Gorman Engel curves for U.S. food consumption is developed and implemented. Box-Cox transformations on prices and income nest functional form. The model nests rank up to rank three. The model is estimated by nonlinear three-stage least squares with annual time series data on 21 foods, 17 nutrients, age and race demographics, and the distribution of income for 1919-1941 and 1947-2000. Results are consistent with full rank three. Point estimates for the Box-Cox parameters on income and prices are 0.86 and 1.09, respectively, strongly rejecting zero and one in both cases. No statistical evidence of serial correlation, specification errors, or parameter instability is found.

Key Words: Aggregation, food demand, functional form, parameter stability, rank, specification errors

JEL Classification: D12, E21

†Jeffrey T. LaFrance is a professor in the School of Economic Sciences at Washington State University, 101 Hulbert Hall, Pullman, WA 99164-6210, telephone 509-335-8512, fax 509-335-1173, email jtlafrance@wsu.edu. I am grateful to the USDA/ERS for their support of this research under cooperative agreement # 43-3AEM-2-80072.

The Structure of U.S. Food Demand

1. Introduction

Over the past several years, farm and food policy in the United States has undergone a continuous transformation. The 1996 farm bill replaced farm-level price and income supports with decoupled payments. Welfare, Food Stamps, Women, Infants and Children (WIC), Aid to Families with Dependent Children (AFDC), and School Lunch programs also were reduced in scope and replaced with block grants. Federally subsidized crop insurance has increased from 28 crops on 26 million acres with total liability of $6 billion in 1980 to more than 100 crops on over 220 million acres with total liability in excess of $46 billion in 2006. A current proposal is to expand this program to all forage on public and private grazing land – a net addition of more than 800 million acres. Subsidies for corn ethanol were $7 billion in 2006, leading to significant diversions of wheat, soybean, and other crop acres to corn production in 2007, and rapidly increasing prices for crops and the foods they produce.

These policies and programs and changes in them all influence retail food prices, food quantities consumed, nutrition, food expenditures, and the net incomes of consumers and taxpayers. But we understand poorly their joint effects on the economic well being, food consumption, and nutrition for U.S. consumers. This calls for a coherent, internally consistent model of the demand for food and nutrition. Developing and implementing one such model is the focus of this paper.

As one motivating example among many possible alternatives, consider the joint incentive effects of food stamps and the dairy program. Food stamps provide subsidies for food consumption in an effort to increase the nutritional status of the poor. In contrast, price discrimination in milk marketing orders increase prices paid for fresh milk and lower prices paid for manufactured dairy products (Heien, 1978; Ippolito and Masson, 1978; LaFrance and de Gorter, 1985). These relative price effects cause households to substitute away from fresh milk and towards processed dairy products. Nutritionists and healthcare professionals argue that processed foods containing relatively high levels of fat, cholesterol, salt, sugar, and additives are considerably less healthy than fresh foods that do not contain these factors and are high in fiber, vitamins, and minerals. The upshot is that many farm-level price and income support programs and policies create incentives in direct opposition to those created by food subsidy programs targeting consumers.

The next section of the paper analyzes the theoretical and econometric issues associated with modeling U.S. food consumption and the implied demand for nutrients. Section three characterizes the econometric model and its properties. Section four discusses the data, empirical results, hypothesis tests, and model diagnostics. The fifth section summarizes and concludes.

2. Modeling Food Demand

A central focus of a great deal of research on farm and food policy and consumer choice has been an attempt to forge the links between food consumption choices and nutrition.

Almost all of the research in this area uses aggregate data, although limited attention has been paid to the implications of aggregation on the structure of economically consistent demand systems. A limited but important and influential subset of the literature on the theory of exactly aggregable demand systems includes: Gorman (1953, 1961, 1965, 1981); Muellbauer (1975, 1976); Deaton (1975, 1986); Howe, Pollak and Wales (1979); Deaton and Muellbauer (1980); Jorgenson, Lau and Stoker (1980, 1982); Russell (1983, 1996); Jorgenson and Slesnick (1984, 1987); Lewbel (1987a, 1988;, 1989a, 1989b, 1990, 1991, 2003, 2004); Diewert and Wales (1987, 1988); Blundell (1988a, 1988b); Wales and Woodland (1983); Brown and Walker (1989); van Daal and Merkies (1989); Jerison (1993); Russell and Farris (1993, 1998); Stoker (1993); Banks, Blundell, and Lewbel (1997); LaFrance, Beatty, Pope and Agnew (2000, 2002); LaFrance (2004); and La-France, Beatty, and Pope (2005, 2006).[1]

In this section, I present a method to identify and estimate the impacts of policies on food consumption, nutrition, and consumer welfare using market data by developing and analyzing an exactly aggregable model of U.S. food demand. The purposes of this approach are to derive and implement a theoretically consistent empirical model of household food consumption which: (1) nests the number (i.e., the *rank* of the demand system) and the functional form of the income terms, and the functional form of the price terms in food demand; (2) admits consistent, asymptotically efficient estimation of the food demand parameters with aggregate data; and (3) allows us to draw inferences on the nutritional and welfare effects of farm and food policies on consumers – both in the aggregate and for specific income and other demographic categories.

We first require a fairly large amount of notation. Let $\boldsymbol{p} \in \mathbb{R}_{++}^{n_q}$ be the n_q–vector of market prices for foods, $\boldsymbol{q} \in \mathbb{R}_{+}^{n_q}$, let $\tilde{\boldsymbol{p}} \in \mathbb{R}_{++}^{n_{\tilde{q}}}$ be the $n_{\tilde{q}}$–vector of market prices for other goods, $\tilde{\boldsymbol{q}} \in \mathbb{R}_{+}^{n_{\tilde{q}}}$, let $m \in \mathbb{R}_{++}$ be income, let $s = \tilde{\boldsymbol{p}}^\intercal \tilde{\boldsymbol{q}} = m - \boldsymbol{p}^\intercal \boldsymbol{q} > 0$ be expenditure on other goods, let $\boldsymbol{z} \in \mathbb{R}^K$ be a vector of demographic variables and other demand shifters, let $\pi(\tilde{\boldsymbol{p}})$ be a known, positive-valued, 1° homogeneous, increasing, and weakly concave function of other goods prices, let $\boldsymbol{x} = \left[g_1(p_1 / \pi(\tilde{\boldsymbol{p}})) \quad \dots \quad g_n(p_n / \pi(\tilde{\boldsymbol{p}})) \right]^\intercal \equiv \boldsymbol{g}(\boldsymbol{p} / \pi(\tilde{\boldsymbol{p}}))$ be a vector of twice continuously differentiable, strictly increasing functions of deflated prices of foods, and let $y = f(m / \pi(\tilde{\boldsymbol{p}}))$ be a twice continuously differentiable, strictly increasing transformation of deflated income.

Second, we will make extensive use of the real-valued functions,

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}^\intercal \boldsymbol{B} \boldsymbol{x} + 2\boldsymbol{\gamma}^\intercal \boldsymbol{x} + 1 , \tag{1}$$

$$\theta(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \boldsymbol{z}) = \alpha_0(\tilde{\boldsymbol{p}}, \boldsymbol{z}) + \boldsymbol{\alpha}(\tilde{\boldsymbol{p}}, \boldsymbol{z})^\intercal \boldsymbol{x} , \tag{2}$$

[1] LaFrance and Pope (2008) synthesize this class of models and extend it to full rank rational demand systems of arbitrary rank in a manner that admits nesting and testing for aggregation, rank, functional form, flexibility, and global regularity of the system of demand equations. Appendix A of the expanded version of this paper discusses the extension of this class of models to incomplete demand systems.

where $\boldsymbol{\alpha}(\tilde{\boldsymbol{p}},\boldsymbol{z})$ is a vector of functions of other prices and demographics, $\alpha_0(\tilde{\boldsymbol{p}},\boldsymbol{z})$ is a scalar function of other prices and demographics (Pollak and Wales, 1981), both $\boldsymbol{\alpha}(\tilde{\boldsymbol{p}},\boldsymbol{z})$ and $\alpha_0(\tilde{\boldsymbol{p}},\boldsymbol{z})$ are 0° homogeneous in $\tilde{\boldsymbol{p}}$, $\boldsymbol{B}$ is an $n_q \times n_q$ matrix of parameters, and $\boldsymbol{\gamma}$ is an n_q-vector of parameters. Due to 0° homogeneity of $\boldsymbol{\alpha}(\tilde{\boldsymbol{p}},\boldsymbol{z})$ and $\alpha_0(\tilde{\boldsymbol{p}},\boldsymbol{z})$ in $\tilde{\boldsymbol{p}}$, without any loss in generality, we can (and from this point forward, do) assume that $(\boldsymbol{p},\tilde{\boldsymbol{p}},m)$ are deflated by $\pi(\tilde{\boldsymbol{p}})$. From this point forward, I abuse notation slightly and absorb the deflator into the symbols for the price and income variables, so that $(\boldsymbol{p},\tilde{\boldsymbol{p}},m)$ denote deflated prices and income.

The starting point for the econometric model of U.S. food demand is the class of full rank three exactly aggregable indirect utility functions derived in LaFrance, Beatty, and Pope (2005, 2006) and defined by

$$v(\boldsymbol{x},\tilde{\boldsymbol{p}},y,\boldsymbol{z}) = \psi\left\{-\frac{\sqrt{\varphi(\boldsymbol{x})}}{\left[y-\theta(\boldsymbol{x},\tilde{\boldsymbol{p}},\boldsymbol{z})\right]} - \frac{\boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{x}}{\sqrt{\varphi(\boldsymbol{x},\tilde{\boldsymbol{p}})}}, \tilde{\boldsymbol{p}}, \boldsymbol{z}\right\}. \tag{3}$$

Useful choices for the functions $f(\cdot)$ and $\boldsymbol{g}(\cdot)$ are translated Box-Cox transformations, $y = (m^{\kappa}-1+\kappa)/\kappa$ and $x_i = (p_i^{\lambda}-1+\lambda)/\lambda$, $i = 1,\cdots,n_q$. Note that $\kappa = 1$ implies that $y = m$ and $\kappa = 0$ implies that $y = 1+\ln m$. Analogous relations apply to λ. Thus, this choice produces a demand system that nests the extended price independent generalized linear (PIGL) and the price independent generalized logarithmic (PIGLOG) functional forms (Muellbauer, 1975, 1976). That is, if $\kappa = \lambda = 0$, then we have a full rank three extended translog model (Christensen, Jorgenson and Lau, 1975), while if $\kappa = \lambda = 1$, then we have a full rank three extended quadratic expenditure system (Howe, Pollak, and Wales, 1979; van Daal and Merkies, 1989). For all values of (κ,λ), we obtain a full rank three quadratic price independent generalized linear (QPIGL) or price independent generalized logarithmic (QPIGLOG) demand system. I call this the *generalized quadratic price independent generalized linear incomplete demand system* (GQ-PIGL-IDS).

A full rank two version of the demand system results when $\boldsymbol{\delta} = \mathbf{0}$, while if $\theta \equiv 0$ and $\boldsymbol{\delta} = \mathbf{0}$ then we have a rank one (homothetic) version. Thus, we are able to simultaneously nest rank and functional form of the income terms within a single framework.

Applying Roy's identity gives the demand equations for foods as

$$\boldsymbol{q} = m^{1-\kappa}\boldsymbol{P}^{\lambda-1}\left\{\boldsymbol{\alpha} + \left(\frac{y-\theta}{\varphi}\right)(\boldsymbol{Bx}+\boldsymbol{\gamma}) + \left[\frac{\boldsymbol{I}-(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{x}^{\mathsf{T}}}{\varphi}\right]\boldsymbol{\delta}\frac{(y-\theta)^2}{\varphi}\right\}, \tag{4}$$

where $\boldsymbol{P}^{\lambda-1} = \mathbf{diag}\left[p_i^{\lambda-1}\right]$.

This model is nonlinear in income and the demand equations do not aggregate across individuals to per capita income. However, the Gorman class of Engel curves generates theoretically consistent, exactly aggregable models of demand with a small number of statistics concerning the income distribution. In the present case, we need three moments of the income distribution – the cross-section means of $\{m^{1-\kappa}, m, m^{1+\kappa}\}$,

$$\boldsymbol{e} \equiv \boldsymbol{Pq} = \boldsymbol{P}^{\lambda}\left\{\boldsymbol{A}_1(\boldsymbol{x},z)m^{1-\kappa} + \boldsymbol{A}_2(\boldsymbol{x},z)m + \boldsymbol{A}_3(\boldsymbol{x},z)m^{1+\kappa}\right\}, \tag{5}$$

where $\boldsymbol{A}_1(\boldsymbol{x},z) = \boldsymbol{\alpha} - \left(\frac{1+\kappa\theta}{\kappa}\right)(\boldsymbol{Bx}+\boldsymbol{\gamma}) + \left(\frac{1+\kappa\theta}{\kappa}\right)^2\left[\frac{\boldsymbol{I}-(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{x}^{\top}}{\varphi}\right]\boldsymbol{\delta}$,

$\boldsymbol{A}_2(\boldsymbol{x},z) = \left(\frac{\boldsymbol{Bx}+\boldsymbol{\gamma}}{\kappa\varphi}\right) - 2\left(\frac{1+\kappa\theta}{\kappa^2\varphi}\right)[\boldsymbol{I}-(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{x}^{\top}]\boldsymbol{\delta}$, and

$\boldsymbol{A}_3(\boldsymbol{x},z) = \frac{1}{\kappa^2\varphi}[\boldsymbol{I}-(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{x}^{\top}]\boldsymbol{\delta}$.

LaFrance, Beatty, and Pope (2005, 2006) show that $1-\boldsymbol{\delta}^{\top}\boldsymbol{x}(y-\theta)/\varphi > 0$, $y-\theta<0$, $\varphi>0$, and $\boldsymbol{B} = \boldsymbol{LL}^{\top} + \boldsymbol{\gamma\gamma}^{\top}$, where $\boldsymbol{L}$ is lower triangular, are necessary and sufficient for the Slutsky matrix to be symmetric, negative semi-definite in an open neighborhood of $\kappa = \lambda = 1$. The first three conditions are satisfied without imposition for this data set. I impose the system of nonlinear constraints on $\boldsymbol{B} = \boldsymbol{LL}^{\top} + \boldsymbol{\gamma\gamma}^{\top}$ during estimation (Lau, 1978; Diewert and Wales, 1987).

In the empirical application reported below, the estimated lower triangular matrix $\boldsymbol{L}$ has a reduced rank due to the fact that the symmetric – but not negative semidefinite – version of the model has an estimated $\boldsymbol{B}$ with seven negative Eigen values. I use the following steps to estimate the curvature restricted model. First, I calculate the Eigen-vector/Eigen-value decomposition $\boldsymbol{B} = \boldsymbol{UDU}^{\top}$ for the symmetric model. Second, I replace the seven negative Eigen values in $\boldsymbol{D}$ with zeroes, say $\tilde{\boldsymbol{D}}$, and find the nearest neighbor (in n_q –dimensional Eigen value space) of the symmetric $\boldsymbol{B}$ that is positive semidefinite, say $\tilde{\boldsymbol{B}} = \boldsymbol{U}\tilde{\boldsymbol{D}}\boldsymbol{U}^{\top}$. Third, I construct the lower triangular Choleski factorization $\tilde{\boldsymbol{B}} = \boldsymbol{LL}^{\top}$. In this Choleski factor, $\boldsymbol{L}$, the last $½ \times 7 \times 8 = 28$ lower triangular elements all vanish, so I fix them as constants.[2] Last, I gradually increase the lower seven main diagonal elements of $\boldsymbol{L}$ to 0.03, which does not impact the nonlinear least squares criterion in any observable fashion to six significant digits, in order to ensure that the restricted matrix $\hat{\boldsymbol{B}}$ is symmetric, negative definite. This generates a 22×22 matrix

[2] In analyzing curvature restrictions with this and other data sets, and for other model specifications, these steps for obtaining starting values lead to convergence in a relatively short amount of computational time for this type of highly nonlinear curvature restricted demand model. In the present case, I use a conditional grid search and simulated annealing to confirm that the estimates are a global minimum of the nonlinear least squares criterion. Also note that one can use either an upper or a lower triangular $\boldsymbol{L}$ with no loss in generality in this sequence of steps.

$$\begin{bmatrix} (\hat{\boldsymbol{L}}\hat{\boldsymbol{L}}^{\top} + \hat{\gamma}\hat{\gamma}^{\top}) & \hat{\gamma} \\ \hat{\gamma}^{\top} & 1 \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{L}} & \hat{\gamma} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{L}}^{\top} & \mathbf{0} \\ \hat{\gamma}^{\top} & 1 \end{bmatrix}$$

with positive Eigen values, the smallest of which is 5.93×10^{-7}. These steps produce an estimated Slutsky matrix that is negative semidefinite at all data points.[3]

2.1 Estimating the U.S. Income Distribution

The U.S. Bureau of the Census publishes annual quintile ranges, intra-quintile means, the top five-percentile lower bound, and mean income in the top five-percentile for all U.S. families. This data is available annually beginning in 1947 and in several (although not all) earlier years going back to 1910. In previous work, LaFrance, Beatty, Pope and Agnew (2000, 2002) and Beatty and LaFrance (2005) estimated a system of regression equations to predict these missing data early in the 20th century, and constructed several continuous annual income distributions for all U.S. families, including the log-normal,

$$f_{M_F}(m;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma m} \exp\left\{ -\frac{1}{2\sigma^2} [\ln(m) - \mu]^2 \right\}. \qquad (6)$$

Their results suggest that the log-normal distribution works well in applied food demand analysis – in fact, better than all other alternatives considered. Their empirical results appear to be quite invariant to the choice for the income distribution.

The log-normal distribution is particularly convenient in the present case. Given year-to-year values for $\{\mu, \sigma\}$, I can calculate a closed-form expression for the moments of the income distribution using the formula

$$E_{M_F}\left(m^{\beta}\right) = \int_{-\infty}^{\infty} e^{\beta(\mu+\sigma z)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = e^{\beta\mu + \beta^2\sigma^2/2}. \qquad (7)$$

Hence, I use the log-normal parameters for the income distribution across families calculated by Beatty and LaFrance (2005). I convert the income distributions for all US families to individuals with the linear change of variables, $m_F = \left(\bar{m}_F / \bar{m}_P\right) m_P$, where $\bar{m}_F$ is annual average family income, $\bar{m}_P$ is per capita disposable personal income in that year, which is obtained from annual issues of the *Economic Report of the President*, and m_F and m_P are annual incomes for an individual family or person, respectively. This generates a log-normal income distribution across individuals with the density function,

$$f_{M_P}(m;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma m} \exp\left\{ -\frac{1}{2\sigma^2} \left[\ln(m) - \mu - \ln\left(\bar{m}_P / \bar{m}_F\right) \right]^2 \right\}, \qquad (8)$$

i.e., the mean of the log-normal distribution changes from μ to $\mu + \ln\left(\bar{m}_P / \bar{m}_F\right)$ and the

[3] As a result, the standard errors reported in this paper are potentially biased (Andrews, 1998, 1999).

variance is unaffected.

In the empirical application, mean incomes for families and individuals and the estimated parameters for the income distributions are treated as data, while the parameter κ is estimated jointly with the other parameters in the demand model. Figure 1 presents the time path of the per capita real income distribution parameters and a 3-dimensional plot of the U.S. income distribution for the period 1910-2005.

A large body of empirical evidence exists from cross-sectional studies supporting the hypotheses that the age and ethnic makeup of households have significant effects on consumption choices. To reflect this stylized fact with aggregate US food consumption data, I construct annual measures of the US age and race distribution from several secondary data sources. Age distribution data are available for the entire sample period in eight age intervals: <5; 5-14; 15-24; 25-34; 35-44; 45-54; 55-64; and ≥65 years of age. An analysis of these age data suggests that there are at most four linearly independent categories. Hence, I construct the proportions of the US population in each of the age groups <5, 5-14, 15-24, 25-54 and ≥565, and normalize on the closest approximation to working age adults available, 25-54 years of age since the sum across categories is one in each year.

With respect to the changing distribution of the US population by race, the most refined data available are the percentages of the US population that are White, Black, or Other (neither White nor Black) races. In this case, I normalize on the percent of the population that is White, and include the percentages that are Black or Other as demand shifters along the lines of Pollak and Wales (1981).

Figure 2 presents the time paths of these data for the US population's distribution by race (Figure 2a) and age group (Figure 2b). [4] From the Figure, it is clear that the ethnic makeup of the US population – at least up until the 21st Century – follows a smooth and gradual nonlinear trend from white toward both black and other races. Similarly, the time paths of the US population's age distribution display the birthrate declines during the Great Depression and World War II, with the Baby Boom immediately following the end of the second world war. This trough and bubble ripple through each of the age groups, from youngest to oldest, at predictable time intervals. It is clear from the bottom panel of figure 2 that the upper tail of the US age distribution has increased throughout the century as a nonlinear, smooth and monotonic trend.

[4] I analyzed the data for unit roots. With two exceptions – the percentage of the population that is neither white nor black and the proportion of the population that is ≥55 years old – I did not find strong evidence of integration and found no evidence of cointegration. But with only 77 time series observations, tests for integration and cointegration have very low power. The explanatory variables are quite correlated, as is generally the case when one uses aggregate time series data. This has a significant and deleterious effect on the statistical precision of the model estimates, especially with respect to the parameters entering $\alpha_0(z_t)$. This problem is similar to α_0 in the Almost Ideal Demand System (Deaton and Muellbauer, 1980).

2.2 Elasticities for the Full Rank 3, GQ–PIGL–IDS

It often is useful to calculate price and income elasticities associated of the demands for food items, $\boldsymbol{q}$. Define the three diagonal matrices $\boldsymbol{P} = \mathbf{diag}[p_i]$, $\boldsymbol{Q} = \mathbf{diag}[q_i]$, and $\boldsymbol{W} = \mathbf{diag}[p_i q_i / m]$. Plugging in the translated Box-Cox transformations of normalized food prices and incomes, the n_q–vector of income elasticities is

$$\boldsymbol{\varepsilon}_m^{\boldsymbol{q}} = m\boldsymbol{Q}^{-1}\frac{\partial \boldsymbol{q}}{\partial m} = (1-\kappa)\boldsymbol{\iota} + \boldsymbol{W}^{-1}\boldsymbol{P}^{\lambda}\left\{\left(\frac{\boldsymbol{Bx}+\boldsymbol{\gamma}}{\varphi}\right) + 2\left[\boldsymbol{I} - \frac{(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{x}^{\top}}{\varphi}\right]\boldsymbol{\delta}\left(\frac{y-\theta}{\varphi}\right)\right\}, \quad (9)$$

where $\boldsymbol{\iota} = [1 \cdots 1]^{\top}$ is an n_q–vector of ones. Similarly, the $n_q \times n_q$ matrix of ordinary price elasticities for $\boldsymbol{q}$ is

$$\begin{aligned} \boldsymbol{E}_{\boldsymbol{p}}^{\boldsymbol{q}} = \boldsymbol{Q}^{-1}\frac{\partial \boldsymbol{q}}{\partial \boldsymbol{p}^{\top}}\boldsymbol{P} = (\lambda-1)\boldsymbol{I} + m^{-\kappa}\boldsymbol{W}^{-1}\boldsymbol{P}^{\lambda}&\left\{-\frac{(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{\alpha}^{\top}}{\varphi}\right. \\ +\left[\boldsymbol{B} - 2\frac{(\boldsymbol{Bx}+\boldsymbol{\gamma})(\boldsymbol{Bx}+\boldsymbol{\gamma})^{\top}}{\varphi} - 2\left(\boldsymbol{I} - \frac{(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{x}^{\top}}{\varphi}\right)\boldsymbol{\delta\alpha}^{\top}\right]&\left(\frac{y-\theta}{\varphi}\right) \quad (10) \\ \left.-\left[(\boldsymbol{Bx}+\boldsymbol{\gamma})\boldsymbol{\delta}^{\top} + \boldsymbol{\delta}^{\top}\boldsymbol{x}\left(\boldsymbol{B} - 4\frac{(\boldsymbol{Bx}+\boldsymbol{\gamma})(\boldsymbol{Bx}+\boldsymbol{\gamma})^{\top}}{\varphi}\right)\right]\left(\frac{y-\theta}{\varphi}\right)^2\right\}&\boldsymbol{P}^{\lambda}. \end{aligned}$$

2.3 The Derived Demand for Nutrition

Nutritional intake can be thought of as the result of a production process that uses foods as inputs. The total amount of nutrients consumed is a linear function of the amount of food ingested. Thus, let $\boldsymbol{N}$ denote the $n_z \times n_q$ matrix of nutrient content per unit of food. The ij^{th} entry represents the amount of nutrient i per unit of food j. Given information on (or an estimate of) $\boldsymbol{N}$, we can analyze the policy effects on nutritional intakes using the previously described demand model, since

$$\boldsymbol{h} = \boldsymbol{Nq}\,, \quad (11)$$

where $\boldsymbol{h}$ is an n_h–vector of nutrients important to the household (Lancaster, 1966, 1971; Michael and Becker, 1973; Muth, 1966).

Linearity is justified on the following grounds. If one eats two (identical) eggs, then twice as much protein and twice as much cholesterol is obtained relative to one egg. This proportionality applies to any combination of foods and nutrients. Hence, the relationship between food consumed and nutrient intake is 1° homogeneous. Moreover, if one eats an egg and a steak, the total protein (cholesterol) obtained is the sum of the protein (cholesterol) from the egg plus the protein (cholesterol) from the steak. Hence, the relationship between food consumed and nutrient intake is additive. As in the case of homogeneity, additivity applies to any combination of foods and nutrients. These two properties define linearity; hence, (11) is an identity showing precisely how nutritionists and other public health professionals calculate nutrient intake from food consumption.

An important element in inferring the effect of policies on nutrient intake is to calculate both the ordinary and uncompensated price elasticities of nutrient *i*. It is straightforward to show that these elasticities satisfy

$$\varepsilon_{p_k}^{h_i} = \sum_{j=1}^{n_q} s_{ij} \varepsilon_{p_k}^{q_j} , \tag{12}$$

where, $\varepsilon_{p_k}^{h_i} \equiv (p_k / h_i) \cdot (\partial h_i / \partial p_k)$ is the (ordinary or compensated) price elasticity of nutrient *i* with respect to the price of food *k*, $\varepsilon_{p_k}^{q_j} \equiv (p_k / q_j) \cdot (\partial q_j / \partial p_k)$ is the corresponding (ordinary or compensated) price elasticity of food *j* with respect to the price of food *k*, and $s_{ij} \equiv n_{ij} q_j / h_i$ is the proportion of nutrient *i* that is contributed by food item *j*. Similarly, the income elasticity of nutrient *i* satisfies

$$\varepsilon_{m}^{h_i} = \sum_{j=1}^{n_q} s_{ij} \varepsilon_{m}^{q_j} , \tag{13}$$

where $\varepsilon_{m}^{q_j} \equiv (m / q_j) \cdot (\partial q_j / \partial m)$ is the income elasticity of food *j*. The empirical results and data set compiled for this study can be used to estimate both $[\varepsilon_{p_k}^{h_i}]$ and $[\varepsilon_{m}^{h_i}]$. Because the demand equations are exactly aggregable, so are the implied demands for nutrients by the linearity of (11), as are the price and income elasticities for nutrient intakes. In addition, it is easy to see that, so long as the column rank of $\boldsymbol{N}$ is at least 3, the rank of the derived demands for nutrients is the same as that of the demands for foods.[5]

3. Data

The data set consists of annual time series observations for the period 1919-2000 for all variables. However, I persistently found a significant structural break corresponding to World Ward II for the years 1942-1946. Hence, throughout the paper, the sample period is 1919-1941 plus 1947-2000. Annual observations on the per capita consumption of twenty-one food items can be broken down into four general categories: (1) *dairy products* – milk, butter, cheese, ice cream and frozen yogurt, and canned and powdered milk; (2) *meats, fish, and poultry* – beef, pork, lamb and other red meat, fish and shellfish, and poultry; (3) *fruits and vegetables* – fresh citrus fruit, fresh non-citrus fruit, fresh vegetables excluding potatoes, potatoes, processed fruit, and processed vegetables; and (4) *miscellaneous foods* – margarine and cooking fat and oil, eggs, pasta and cereal grains, sugar and caloric sweeteners, and coffee, tea and cocoa. Average retail prices for these foods

[5] The expanded version of this paper, available at http://www.ses.wsu.edu/People/lafrance.html, contains a complete set of time series data and graphs for the ordinary and compensated price elasticities and the income elasticities for this demand model 1919-41 and 1947-2000, including those for nutrients.

were constructed from several USDA and BLS sources. The consumer price index (CPI) for all nonfood items is used to measure the cost of nonfood expenditures and as the common deflator in the demand equations.

With the generous assistance of the USDA's Human Nutrition Information Service (HNIS), I also compiled annual estimates of the availability of the following seventeen nutrients: (1) *macronutrients* – protein, energy, fat, carbohydrates, and cholesterol; (2) *vitamins* – A, B_6, B_{12}, C, niacin, riboflavin, and thiamin; and (3) *minerals* – calcium, iron, magnesium, phosphorous, and zinc, for each of the twenty-one foods. As outlined in the previous section, the relationship $\boldsymbol{h} = \boldsymbol{N}\boldsymbol{q}$ links prices, incomes, demographics, and government policies to the demand for nutrition derived from the demand for food.

It is by now well-understood that prices and quantities are jointly determined, and that disposable income is measured with error and endogenous to the life cycle consumption decisions of consumers (Attfield, 1985, 1991; Blundell, 1988a, 1988b). Addressing this during the estimation process requires a set of valid instruments for retail food prices and for the moments of the income distribution. The instruments that I use for each retail food price include: a constant term; the demographic variables (proportions of the population that are black and that are neither white nor black, and proportions of the population that are in age groups less than 5 years, 5-14 years, 15-24 years, and 55 years or older); real prices received by farmers in the most recent past marketing year – including the value of government payments – for 22 commodities (wheat, barley, corn, oats, rice, rye, flax, beans, peanuts, potatoes, sweet potatoes, cotton, hay, beef calves, beef cattle, hogs, lambs, sheep, milk cows, milk, eggs, and broilers), the unemployment rate, the real rate of return on 30-year AAA corporate bonds, the manufacturing wage rate, producer price indices for materials and components and for fuels, energy and power, a post-World War II dummy equal to 0 in 1919-1941 and 1 in 1947-2000; and the lagged food price. The unemployment rate is treated as a static shock with a common factor in all of the dynamic first-stage regression equations (Harvey, 1990, 1993; Hendry, 1995).

For the income distribution, a useful instrument is the predicted value of per capita disposable personal income. The first stage instrument for the moments of the income distribution is the predicted value from a time series regression equation for per capita disposable personal income with a constant term, the above list of demographic variables, the unemployment rate, the rate of return to 30-year corporate bonds, the manufacturing wage rate, producer price indices for materials and components and for fuels, energy and power, the post-World War II dummy; and lagged real income. As in the first-stage price equations, the unemployment rate is treated as a static shock with a common factor in the dynamic regression equation.

As is common in NL3SLS and GMM estimation, I use the same instruments in all of the empirical demand equations: the predicted value from each first-stage price equation is an instrumental variable for that price; the predicted value of per capita personal disposable income is an instrument for the moments of the income distribution; and a constant term plus the age and race variables listed above. Under standard assumptions on the weak exogeneity of the instruments (Engel, Hendry, and Richard, 1983), these are known give consistent, asymptotically normal parameter estimates (Amemiya, 1985; Hansen, 1982; Malinvaud, 1980).

The error terms in demand equations are likely to be heteroskedastic (Brown and Walker, 1989), so I estimated standard errors that are robust to heteroskedasticity of an unknown form (White, 1980). However, this correction does not appear to affect the model inferences in this data set. Hence, I only report classical standard errors here. I also tested for serial correlation in the errors, but found no evidence of this in this data set.

4. Specification Errors and Parameter Stability Tests

As stated in the previous section, the original sample period for the empirical application is19-2000.[6] This period includes the Great Depression, World War II, the OPEC Oil Embargo, and Iran-Iraq War. *Ex post*, it stretches the imagination to suppose that the structure of U.S. food demand remained constant throughout this period. On the other hand, this is an interesting empirical question, especially given recent empirical work on structural change in the demand for food and individual food groups.

Many diagnostic procedures for testing parameter stability and model specification errors have been developed (e.g., Brown, Durbin and Evans, 1975; Ploberger and Krämer, 1992). Few of these are designed for large systems of nonlinear simultaneous equation systems with a small sample size. In the present case, the data set provides roughly five degrees of freedom per estimated structural parameter. This precludes the use of recursive-forecast residuals or Chow tests based on sequential sample splits to analyze model specification and/or parameter stability. Even so, it is desirable to have at least some reasonable idea of the degree to which the data are consistent with the model's specification and the hypothesis of constant parameters over time. In this section, I present a set of model specification and parameter stability diagnostic statistics using in-sample estimated residuals. These statistics have power against a range of local alternatives, including nonstationary parameters, model specification errors, or the restrictions on preference heterogeneity that are required for exact aggregation.

The main ideas are simple and straightforward. If the model is stationary and the errors are innovations, then consistent estimates of the model parameters can be found in any number of ways. Given consistent parameter estimates, the estimated errors converge in probability, hence in distribution, to the true errors, $\hat{\boldsymbol{\varepsilon}}_t \xrightarrow{P} \boldsymbol{\varepsilon}_t$. For each $i = 1, \cdots, n_q$, by the central limit theorem for stationary Martingale differences, we have

$$\frac{1}{\sqrt{T}\sigma_i} \sum_{t=1}^{T} \varepsilon_{it} \xrightarrow{D} N(0,1), \tag{14}$$

[6] In this study, and in substantial previous work with this data, I have found consistent strong statistical evidence of a structural break during the World War II period using the methods developed in this section. Hence, the sample period used throughout this paper is, 1919-1941 and, 1947-2000. The test results I report below are for this sample period.

where $\sigma_i^2 = E(\varepsilon_{it}^2)$ is the variance of the residual for the i^{th} demand equation. If we use a fixed proportion of the sample to construct a partial sum of the model's error terms, then we have, uniformly in $z \in [0,1]$,

$$\frac{1}{\sqrt{T}\sigma_i}\sum_{t=1}^{[zT]}\varepsilon_{it} \xrightarrow{D} N(0,z), \tag{15}$$

where $[zT]$ is the largest integer that does not exceed zT. The variance is z because we are summing $[zT]$ independent terms each with variance $1/T$. Multiplying (14) by z and subtracting from (15) then gives

$$\frac{1}{\sqrt{T}\sigma_i}\sum_{t=1}^{[zT]}\left(\varepsilon_{it} - \overline{\varepsilon}_i\right) \xrightarrow{D} W(z) - zW(1) \equiv B(z), \tag{16}$$

where $W(z)$ is a standard Brownian motion on the unit interval, with $W(z) \sim N(0,z)$, and $B(z)$ is a standard Brownian bridge, or tied Brownian motion. For all $z \in [0,1]$, $B(z)$ has an asymptotic Gaussian distribution, with mean zero and standard deviation $\sqrt{z(1-z)}$ (Bhattacharya and Waymire, 1990). For a given z – that is, to test for a break point in the model at a fixed and known date – an asymptotic 95% confidence interval for $B(z)$ is $\pm 1.96\sqrt{z(1-z)}$. To check for an unknown break point, the statistic

$$Q_T = \sup_{z \in [0,1]} \left|B_T(z)\right| \tag{17}$$

has an asymptotic 5% critical value of 1.36 (Ploberger and Krämer, 1992).

We can use consistently estimated residuals and consistently estimated standard errors to obtain sample analogues to these asymptotic Brownian bridges. This gives

$$B_{iT}(z) \equiv \frac{1}{\sqrt{T}\hat{\sigma}_i}\sum_{t=1}^{[zT]}(\hat{\varepsilon}_{it} - \overline{\hat{\varepsilon}}_i) \xrightarrow{D} B(z), \tag{18}$$

also uniformly in $z \in [0,1]$, so long as the model specification is correct and the parameters are constant across time periods. This statistic is a single equation first-order specification/parameter stability statistic since it is based on the first-order moment conditions, $E(\varepsilon_{it}) = 0 \ \forall \ i,t$.[7] A system-wide first-order specification/parameter stability statistic can be defined by

---

[7] Subtracting the residual sample means is innocuous when each equation includes a free intercept term. Integrable demand systems generally do not include independent intercepts due to Slutsky symmetry. In such a case, the statistics $\sqrt{T}\,\overline{\hat{\varepsilon}}_i / \hat{\sigma}_i \xrightarrow{D} N(0,1)$ can be used to check the first-order moment condition of each demand equation. Both single equation sample mean and Brownian bridge statistics are reported below.

$$B_T(z) \equiv \frac{1}{\sqrt{T}} \sum_{t=1}^{[zT]} \left[\frac{1}{\sqrt{n_q}} \sum_{i=1}^{n_q} (\hat{\xi}_{it} - \bar{\hat{\xi}}) \right] \xrightarrow{D} B(z), \tag{19}$$

where $\hat{\boldsymbol{\xi}}_t = \hat{\boldsymbol{\Sigma}}^{-1/2} \hat{\boldsymbol{\varepsilon}}_t$ is the t^{th} estimated standardized error vector and $\bar{\hat{\xi}} \equiv \sum_{t=1}^{T} \sum_{i=1}^{n_q} \hat{\xi}_{it} / n_q T$.

Similar methods can be applied to test for second-order nonstationarity. I focus on the system wide test statistic. Let $\boldsymbol{\Sigma}$ be factored into $\boldsymbol{L}\boldsymbol{L}^{\top}$, where $\boldsymbol{L}$ is lower triangular and nonsingular. Define the random vector $\boldsymbol{\xi}_t$ by $\boldsymbol{\varepsilon}_t = \boldsymbol{L}\boldsymbol{\xi}_t$. In addition to the assumptions above, add $\sup_{i,t} E(\varepsilon_{it}^4) < \infty$. Estimate the within-period average sum of squared standardized residuals by

$$\hat{\upsilon}_t = \tfrac{1}{n_q} \hat{\boldsymbol{\xi}}_t^{\top} \hat{\boldsymbol{\xi}}_t = \tfrac{1}{n_q} \hat{\boldsymbol{\varepsilon}}_t^{\top} \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\varepsilon}}_t, \tag{20}$$

where $\hat{\boldsymbol{\varepsilon}}_t$ is the vector of consistently estimated residuals in period t and $\hat{\boldsymbol{\Sigma}} = \sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t^{\top} / T$ is the associated consistently estimated error covariance matrix. The mean of the true υ_t is one for each t, and the martingale difference property of $\boldsymbol{\varepsilon}_t$ is inherited by $\upsilon_t - 1$.[8] A consistent estimator of the asymptotic variance of υ_t is

$$\hat{\sigma}_{\upsilon}^2 = \frac{1}{T} \sum_{t=1}^{T} (\hat{\upsilon}_t^2 - 1). \tag{21}$$

A system wide second-order specification/parameter stability test is obtained by calculating centered and standardized partial sums of $\hat{\upsilon}_t$,

$$B_T(z) = \frac{1}{\sqrt{T}\hat{\sigma}_{\upsilon}} \cdot \sum_{t=1}^{[zT]} (\hat{\upsilon}_t - 1) \xrightarrow[T \to \infty]{D} B(z), \tag{22}$$

uniformly in $z \in [0,1]$, where the limiting distribution on the far right follows from the identity $\bar{\hat{\upsilon}} \equiv \sum_{t=1}^{T} \hat{\upsilon}_t / T \equiv 1$.

5. Empirical Results

The empirical model has per capita real food expenditures as the dependent variables and specifies linear functions for $\alpha_0(\boldsymbol{z}_t)$ and $\boldsymbol{\alpha}(\boldsymbol{z}_t)$,

$$\boldsymbol{e}_t = \boldsymbol{P}_t^{\lambda} \left[\boldsymbol{A}_1(\boldsymbol{x}_t, \boldsymbol{z}_t) E_M(m_t^{1-\kappa}) + \boldsymbol{A}_2(\boldsymbol{x}_t, \boldsymbol{z}_t) E_M(m_t) + \boldsymbol{A}_3(\boldsymbol{x}_t, \boldsymbol{z}_t) E_M(m_t^{1+\kappa}) \right] + \boldsymbol{\varepsilon}_t, \tag{1.23}$$

[8] In particular, $\bar{\hat{\upsilon}} = \sum_{t=1}^{T} \hat{\upsilon}_t / T = \sum_{t=1}^{T} \hat{\xi}_t^{\top} \hat{\xi}_t / n_q T = \operatorname{tr}\left[\left(\sum_{\tau=1}^{T} \hat{\boldsymbol{\varepsilon}}_\tau \hat{\boldsymbol{\varepsilon}}_\tau^{\top} / T \right)^{-1} \left(\sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t^{\top} / T \right) / n_q \right] = 1.$

where $\boldsymbol{A}_1(\boldsymbol{x}_t, \boldsymbol{z}_t) = \boldsymbol{\alpha}_{10} + \boldsymbol{A}\boldsymbol{z}_t - \left(\frac{1+\kappa\theta_t}{\kappa}\right)(\boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\gamma}) + \left(\frac{1+\kappa\theta_t}{\kappa}\right)^2 \left[\frac{\boldsymbol{I} - (\boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\gamma})\boldsymbol{x}_t^\top}{\varphi_t}\right]\boldsymbol{\delta},$

$$\boldsymbol{A}_2(\boldsymbol{x}_t, \boldsymbol{z}_t) = \left(\frac{\boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\gamma}}{\kappa\varphi_t}\right) - 2\left(\frac{1+\kappa\theta_t}{\kappa^2\varphi_t}\right)[\boldsymbol{I} - (\boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\gamma})\boldsymbol{x}_t^\top]\boldsymbol{\delta},$$

$$\boldsymbol{A}_3(\boldsymbol{x}_t, \boldsymbol{z}_t) = \frac{1}{\kappa^2\varphi_t}[\boldsymbol{I} - (\boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\gamma})\boldsymbol{x}_t^\top]\boldsymbol{\delta},$$

$$\theta_t = \alpha_{00} + \boldsymbol{\alpha}_{01}^\top \boldsymbol{z}_t + (\boldsymbol{\alpha}_{10} + \boldsymbol{A}\boldsymbol{z}_t)^\top \boldsymbol{x}_t,$$

$$\varphi_t = \boldsymbol{x}_t^\top \boldsymbol{B}\boldsymbol{x}_t + 2\boldsymbol{\gamma}^\top \boldsymbol{x}_t + 1,$$

$$\boldsymbol{x}_t = \left[(p_{1t}^\lambda + \lambda - 1)/\lambda \ \cdots \ (p_{n_q t}^\lambda + \lambda - 1)/\lambda\right]^\top,$$

$$\boldsymbol{z}_t = \left[age_{<5,t}\ age_{5-14,t},\ age_{15-24,}\ age_{\geq 55,t},\ black_t,\ other_t\right]^\top,$$

and $\boldsymbol{\varepsilon}_t$ $i.i.d.$ $(\boldsymbol{0}, \boldsymbol{\Sigma})$ $\forall$ t.[9]

This demand model is estimated with nonlinear three-stage least squares (NL3SLS). The high degree of nonlinearity in the parameters, especially the Box-Cox exponents, implies that one needs to carefully check whether a global minimum of the generalized residual sum of squares criterion has been found. Figure 3 presents the results of an effort to conduct a comprehensive analysis of this issue. I condition the second-stage NL3SLS objective function on (κ, λ) and the estimated error covariance matrix $\hat{\boldsymbol{\Sigma}}$ and search over $(\kappa, \lambda) \in [0, 1¼] \times [0, 1¼]$. The upper left panel displays the results of this search for an open neighborhood of the optimal point estimates. The objective function is well-behaved throughout this neighborhood, although it is clear from the upper right panel that the level curves are not elliptical except in a smaller neighborhood of the optimal point estimates. The NL3SLS criterion is continuous but not differentiable over the full range of parameter values $(\kappa, \lambda) \in [0, 1¼] \times [0, 1¼]$.

I find estimating small values of λ to be most difficult, particularly as κ decreases

[9] I normalize the 22×22 matrix

$$\boldsymbol{B}^* = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{\gamma} \\ \boldsymbol{\gamma}^\top & b_{22,22} \end{bmatrix}$$

so that $b_{22,22} = 1$. This identifying normalization accounts for the fact that the demand system (4) is 0° homogeneous in the elements of $\boldsymbol{B}^*$. Therefore, one and only one element of $\boldsymbol{B}^*$ must be fixed to identify the other model parameters. I normalize on the last diagonal element because nonfood expenditure is >90% of per capita disposable personal income in the sample.

below roughly 0.2. Hence, I condition the NL3SLS criterion on κ only and estimate λ jointly with the other model parameters, while searching over $\kappa \in [0,1¼]$. The results for the conditional sum of squares criterion are depicted in the lower left panel of figure 3, while the conditional estimates for $\lambda(\kappa)$ and the associated conditional standard errors are shown in the lower right panel. These results suggest that I am recovering the global minimum sum of squares in the results presented below.[10]

Table 1 presents the equation summary statistics for the first-stage instrumental variables regression equations. Table 2 contains summary statistics and model diagnostics for the nonlinear three-stage least squares estimates of the structural demand model, which is estimated with deflated food expenditures as dependent variables. In this table, the R^2 is the squared correlation between the observed and predicted dependent variable, $\sqrt{T}\,\bar{\hat{\varepsilon}}_i / \hat{\sigma}_i$ is asymptotically a standard normal under the hypothesis that the means of the error terms are zero, $\hat{B}_{iT}(z) = \sum_{t=1}^{[zT]} (\hat{\varepsilon}_{it} - \bar{\hat{\varepsilon}}_i) / \sqrt{T\hat{\sigma}_{ii}}$ is the single equation asymptotic Brownian bridge test statistic, and the Durbin-Watson statistic is based on the estimated errors from the nonlinear three stage least squares estimates of each demand equation. Based on these criteria, it is notable that there is little evidence of specification errors, nonconstant parameters, or serial correlation in the errors. This is unusual in that the imposition of parameter restrictions such as symmetry and curvature usually introduces substantial serial correlation.

These conclusions carry over to the system-wide hypothesis tests in the same class. The results of these hypothesis tests are .383 and 1.084, respectively, both of which are well within the 95% asymptotic confidence level of 1.36 for a tied Brownian motion (Bhattacharya and Waymire, 1990). A demand systems test for a common autocorrelation coefficient returns an estimate of $\hat{\rho} = .00793$ with an estimated asymptotic standard error of .025 (resulting in a p-value of .752). Asymptotically standard normal test statistics for skewness and excess kurtosis are –.223 and 1.374, respectively, yielding a Jarque-Bera $\chi^2(2)$ test statistic for the null hypothesis that the errors are jointly normal that is equal to 1.937, with an associated p-value of .380. Thus, I conclude that this is a stable and statistically adequate empirical model of aggregate U.S. food demand.

Table 3 presents the estimated parameters for constant terms, demographics, main diagonal elements of $\boldsymbol{B^*}$, Box-Cox parameters (κ,λ), and rank three parameters, $\boldsymbol{\delta}$, for each food demand equation.[11] The main diagonal elements of $\boldsymbol{B^*}$ strongly dominate the re-

---

[10] I search jointly over (κ, λ) throughout the entire rectangle in very small increments and a large number of iterations for each (κ, λ) pair. At times more than 5,000 iterations in TSP© version 4.5 are needed per parameter pair in the lower areas of this rectangle. Each step at times requires several hours on a 3.8GHz Pentium IV workstation with 4GB of RAM. I find no alternative minima and the graph of the NL3SLS surface over this range of values does not lead to any additional insights or information beyond those conveyed here, and looks ragged and unappealing compared to figure 3.

[11] Details in the expanded version of this paper at http://www.ses.wsu.edu/People/lafrance.html, are freely

maining elements of this matrix and few off-diagonal terms are statistically different from zero at any of the usual significance levels.

As should be expected, the high degree of multicollinearity both among the right-hand-side variables in the demand equations and among the instrumental variables, especially the demographics, has a negative effect on the precision of the parameter estimates. For example, in each demand equation, in most cases one or two of the coefficients in the functions $\alpha_i(z_t)$ are statistically significant at the 5% level of better. However, a few equations – beef, eggs, and nonfood expenditures – have three estimated coefficient estimates on the demographic variables and constant terms that are highly significant, and almost all equations have two or more coefficients on these variables that are significant at the 10% level or better.

As a result, it could be that this model is somewhat over-parameterized. For example, one could consider restricting the price coefficients matrix ***B**** to be diagonal, and only including those demographic variables that appear to be statistically significant in the empirical model. However, this kind of modeling approach leads to pre-test biases and is conditional on the chosen sequence of restrictions introduced to simplify the econometric model. Therefore, I only present the empirical results for the general specification here. Moreover, the fact that the individual parameter estimates have wide distributions is useful and informative since this indicates a corresponding degree of uncertainty that one should attach to policy analyses and inferences generated with this data set.

I test the following four hypotheses of interest using Wald tests: (1) functional form: $H_0:\kappa=\lambda=1$, $\chi^2(2)=44.04$, p-value = 2.7×10^{-10}; (2) full rank two, $H_0:\boldsymbol{\delta}=\mathbf{0}$, $\chi^2(21)=67.67$, p-value = 8.3×10^{-7}; (3) the non-separable quadratic utility function, $H_0:\kappa=\lambda=1,\ \boldsymbol{\delta}=\mathbf{0}$, $\chi^2(23)=106.87$, p-value = 9.0×10^{-13}; and (4) the separable quadratic utility function,[12] $H_0:\kappa=\lambda=1,\ \boldsymbol{\gamma}=\boldsymbol{\delta}=\mathbf{0}$, $\chi^2(44)=185.54$, p-value = 0.0. Clearly, the data has the ability to discriminate against these restrictions, the extensions to functional form and rank, and the common assumption of separability of foods from all other goods in consumer preferences.

6. Conclusions

This paper presents results on an econometric model of per capita food consumption and nutritional intake for the United States. The model is consistent with economic theory, exact aggregation, and attains the maximal rank of three with this data set. Parameter restrictions for monotonicity and curvature over an open convex hull of the data are derived and implemented. The empirical application estimates a system of demands for twenty-one food items using annual U.S. per capita time series data for 1919-1941 and 1947-

available for download, along with the complete data set and supplemental Appendices.

[12] It can be shown that this null hypothesis is the *only* way that foods can be weakly separable from all other goods in this class of demand systems.

2000. The results suggest that this empirical model is a reasonable and coherent econometric framework for studying the aggregate consumer effects of changes in farm and food policies in the United States. Ongoing work includes estimating the consumer impacts of US farm and food policies on food demand and nutrition using the empirical results of this model and data set, as well as the distributional effects of taxing the fat and sugar content of foods.

References

Amemyia, T, 1985, *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Andrews, DWK, 1998, "Hypothesis Testing with a Restricted Parameter Space," *Journal of Econometrics*, 84: 155–199.

Andrews, DWK, 1999, "Estimation When a Parameter is on a Boundary," *Econometrica* 67: 1341-1383.

Attfield, CLF, 1985, "Homogeneity and Endogeneity in Systems of Demand Equations." *Journal of Econometrics* 27: 197-209.

Attfield, CLF, 1991, "Estimation and Testing When Explanatory Variables Are Endogenous," *Journal of Econometrics*, 48: 395-408.

Banks, J, R Blundell, and A Lewbel, 1997, "Quadratic Engel Curves and Consumer Demand," *The Review of Economics and Statistics*, 79: 527-539.

Beatty, TKM, and JT LaFrance, "United States Demand for Food and Nutrition in the Twentieth Century." *American Journal of Agricultural Economics* (2005): 1159-1166.

Bhattacharya, RN and EC Waymire, 1990, *Stochastic Processes with Applications*, New York: John Wiley and Sons.

Blundell, R, 1988a, "Econometric Approaches to the Specification of Life-Cycle Labour Supply and Commodity Demand Behavior." *Econometric Reviews* 5: 89-146.

Blundell, R, 1988b, "Consumer Behavior: Theory and Empirical Evidence – A Survey" *The Economic Journal* 98: 16-65.

Brown, RL, J Durbin, and JM Evans, 1975, "Techniques for Testing the Constancy of Regression Relationships over Time." *Journal of the Royal Statistical Society*, Series B, 37: 149-92.

Brown, BW, and MB Walker, 1989, "The Random Utility Hypothesis and Inference in Demand Systems." *Econometrica* 57: 815-829.

Christensen, LR, DW Jorgensen and LJ Lau, 1975, "Transcendental Logarithmic Utility Functions." *American Economic Review* 65: 367-383.

Deaton, A, 1975, Models and Projections of Demand in Post-War Britain, London: Chapman and Hall.

Deaton, A, 1986, "Demand Analysis," Chapter 30 in the *Handbook of Econometrics*, *Volume III*, Z Griliches and MD Intriligator, eds., New York: Elsevier Science, 1767-1839.

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

Diewert, WE, and TJ Wales, 1987, "Flexible Functional Forms and Global Curvature Conditions." *Econometrica* 55: 43-68.

Diewert, WE, and TJ Wales, 1988, "Normalized Quadratic Systems of Consumer Demand Functions." *Journal of Business and Economic Statistics* 6: 303-312.

Engle, RF, DF Hendry, and JF Richard, 1983, "Exogeneity." *Econometrica* 51: 277-304.

Gorman, WM, 1953, "Community Preference Fields." *Econometrica* 21: 63-80.

Gorman, WM, 1961, "On a Class of Preference Fields." *Metroeconomica* 13: 53-56.

Gorman, WM, 1965, "Consumer Budgets and Price Indices." Unpublished typescript, Reprinted as Chapter 5 in C, Blackorby and T, Shorrocks, eds., *Separability and Aggregation: Collected Works of W, M, Gorman*, Oxford: Clarendon Press, 1995a.

Gorman, WM, 1981, "Some Engel Curves." In A Deaton, ed, *Essays in Honour of Sir Richard Stone*, Cambridge: Cambridge University Press,

Hansen, LP, 1982, "Large Sample Properties of Generalized Methods of Moments Estimators." *Econometrica* 50: 1029-1054.

Harvey, AC, 1990, *The Econometric Analysis of Time Series*, Cambridge, Massachusetts: The MIT Press.

Harvey, AC, 1993, *Time Series Models*, Second Edition, Cambridge, Massachusetts: The MIT Press.

Hendry, DF, 1995, *Dynamic Econometrics* Oxford: Oxford University Press, 1995.

Heien, D, 1977, "The Cost of U.S, Dairy Price Support Programs: 1949-74." *Review of Economics and Statistics* 59: 1-8.

Howe, H, RA Pollak, and TJ Wales, 1979, "Theory and Time Series Estimation of the Quadratic Expenditure System." *Econometrica* 47: 1231-1247.

Ippolito, RA and RT Masson, 1978, "The Social cost of Government Regulation of Milk." *Journal of Law and Economics* 21: 33-65.

Jerison, M, 1993, "Russell on Gorman's Engel Curves: A Correction." *Economics Letters* 23: 171-175,

Jorgenson, DW, LJ Lau, and TM Stoker, 1980, "Welfare Comparisons Under Exact Aggregation." *American Economic Review* 70: 268-272.

Jorgenson, DW, LJ Lau, and TM Stoker, 1982, "The Transcendental Logarithmic Model of Aggregate Consumer Behavior," in *Advances in Econometrics*, R.L, Basmann and G.F, Rhodes, Jr, eds, Greenwich: JAI Press.

Jorgenson, DW and DT Slesnick, 1984, "Aggregate Consumer Behavior and the Measurement of Inequality." *Review of Economic Studies* 51: 369-392.

Jorgenson, DW and DT Slesnick, 1987, "Aggregate Consumer Behavior and Household Equivalence Scales." *Journal of Business and Economic Statistics* 5: 219-232.

LaFrance, JT, 2004, "Integrability of the Linear Approximate Almost Ideal Demand System." *Economic Letters* 84: 297-303.

LaFrance, JT, TKM Beatty, and RD Pope, 2005, "Building Gorman's Nest." *9th World Congress of the Econometrics Society*, London, UK.

LaFrance, JT, TKM Beatty, and RD Pope, 2006, "Gorman Engel Curves for Incomplete Demand Systems." In *Exploring Frontiers in Applied Economics: Essays in Honor of Stanley R, Johnson*, MT Holt and J–P Chavas, ed, Berkeley: BEP.

LaFrance, JT, TKM Beatty, RD Pope and GK Agnew, 2000, "US Income Distribution and Gorman Engel Curves for Food." *Proceedings of IIFET, 200*0.

LaFrance, JT, TKM Beatty, RD Pope and GK Agnew, 2002, "Information Theoretic Measures of the US Income Distribution in Food Demand." *Journal of Econometrics* 107: 235-257.

LaFrance, JT and H de Gorter, 1985, "Regulation in a Dynamic Market: The United States Dairy Industry." *American Journal of Agricultural Economics* 67: 821-832.

LaFrance, JT and RD Pope, 2008, "Full Rank Rational Demand Systems." Department of Agricultural and Resource Economics and Policy, Working Paper #1021 (revised), University of California-Berkeley.

Lancaster, KJ, 1966, "A New Approach to Consumer Theory." *Journal of Political Economy* 74: 132-157.

Lancaster, KJ, 1971, *Consumer Demand: A New Approach*, New York: Columbia University Press.

Lau, LJ, 1978, "Testing and Imposing Monotonicity, Convexity and Quasi-convexity Constraints." Appendix A.4 in *Production Economics: A Dual Approach to Theory and Applications,* M Fuss and D McFadden (eds.), Amsterdam: North-Holland, 409-455.

Lewbel, A, 1987a, "Characterizing Some Gorman systems that satisfy consistent aggregation." *Econometrica* 55: 1451-1459.

Lewbel, A, 1987b, "Fractional Demand Systems." *Journal of Econometrics* 36: 311-337.

Lewbel, A, 1988, "An Exactly Aggregable Trigonometric Engel Curve Demand System." *Econometric Reviews* 2: 97-102.

Lewbel, A, 1989a, "A Demand System Rank Theorem." *Econometrica* 57: 701-705.

Lewbel, A, 1989b, "Nesting the AIDS and Translog Demand Systems." *International Economic Review* 30: 349-356.

Lewbel, A, 1990, “Full Rank Demand Systems.” *International Economic Review* 31: 289-300.

Lewbel, A, 1991, “The Rank of Demand Systems: Theory and Nonparametric Estimation.” *Econometrica* 59: 711-730.

Lewbel, A, 2003, “A Rational Rank Four Demand System.” *Journal of Applied Econometrics* 18: 127-2003, Corrected mimeo, July, 2004.

Malinvaud, E, 1980, *Statistical Methods of Economics*, 2nd Edition, Amsterdam: North-Holland.

Michael, RT and GS Becker, 1973, “On the New Theory of Consumer Behavior.” *Swedish Journal of Economics* 75: 378-396.

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.

Muth, RF, 1966, “Household Production and Consumer Demand Functions.” *Econometrica* 34: 699-708.

Ploberger, W and W Krämer, 1992, “The CUSUM Test with OLS Residuals.” *Econometrica* 60: 271-286.

Pollak, RA and TJ Wales, 1981, “Demographic Variables in Demand Analysis.” *Econometrica* 49: 1533-1551.

Russell, T, 1983, “On a Theorem of Gorman.” *Economic Letters* 11: 223-224.

Russell, T, 1996, “Gorman Demand Systems and Lie Transformation Groups: A Reply.” *Economic Letters* 51: 201-204.

Russell, T and F Farris, 1993, “The Geometric Structure of Some Systems of Demand Functions.” *Journal of Mathematical Economics* 22: 309-325.

Russell, T and F Farris, 1998, “Integrability, Gorman Systems, and the Lie Bracket Structure of the Real Line.” *Journal of Mathematical Economics* 29: 183-209.

Stoker, TM, 1993, “Empirical Approaches to the Problem of Aggregation over Individuals.” *Journal of Economic Literature* 31: 1827-1874.

van Daal, J and AHQM Merkies, 1989, “A Note on the Quadratic Expenditure Model.” *Econometrica* 57: 1439-1443.

Wales, TJ and AD Woodland, 1983, “”Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints.” *Journal of Econometrics* 21: 263-285.

White, H, 1980, “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica* 48: 817-38.

Tables

Table 1. First-Stage Instrumental Variable Regressions Summary Statistics.

	Mean	Standard Deviation	$\hat{\sigma}_u^2$	R^2	Durbin Watson	LM Heter.	Coeff. of Skewness	Excess Kurtosis
Per Capita Income	2600.8	1259.3	1648.3	.9991	1.889	6.016*	0.132	1.012
Fresh Milk & Cream	.1185	.0175	4.403×10^{-6}	.9855	2.449	4.920*	0.195	–0.170
Butter	.7675	.1900	6.675×10^{-4}	.9813	2.072	0.048	0.789*	2.427**
Cheese	.7672	.1353	3.262×10^{-4}	.9819	2.315	0.476	0.308	–0.251
Frozen Dairy	.2038	.0449	2.584×10^{-5}	.9870	2.203	0.596	–.290	1.028*
Canned & Powdered Milk	.1699	.0234	1.559×10^{-5}	.9712	2.327	0.015	0.377	0.703
Beef	.9739	.2195	1.425×10^{-3}	.9700	1.990	0.109	0.189	0.515
Pork	.6097	.1184	8.899×10^{-4}	.9357	1.864	5.22×10^{-3}	0.464	1.583**
Other Red Meat	.7470	.1299	6.608×10^{-4}	.9603	1.986	3.87×10^{-3}	0.697*	0.582
Fish	.6881	.2218	2.4637×10^{-4}	.9949	1.931	6.978*	0.087	–0.342
Poultry	.4766	.1873	2.745×10^{-4}	.9921	2.386	0.167	0.750**	0.931
Fresh Citrus Fruit	.1607	.0434	2.126×10^{-4}	.8855	2.396	12.83**	0.951**	2.407**
Fresh Non-citrus Fruit	.1661	.0266	3.937×10^{-5}	.9438	1.801	1.533	0.086	0.943
Fresh Vegetables	.1834	.0317	5.234×10^{-5}	.9470	2.184	4.93×10^{-4}	0.354	1.095*
Potatoes	.0713	.0178	2.693×10^{-5}	.9136	2.143	3.32×10^{-4}	0.123	–0.121
Processed Fruit	.5099	.0595	3.770×10^{-4}	.8921	2.425	0.1199	0.110	–0.574
Processed Vegetables	.1945	.0235	1.711×10^{-5}	.9687	2.206	0.0172	0.797**	1.236**
Margarine, Fats & Oils	.3244	.0964	1.573×10^{-4}	.9829	1.893	4.74×10^{-5}	0.047	1.030
Eggs	.3753	.1571	2.989×10^{-4}	.9877	2.379	1.008	0.248	0.434
Pasta & Cereal Grains	.1168	.0134	3.686×10^{-6}	.9794	1.934	1.250	–.103	0.223
Caloric Sweeteners	.2227	.0440	7.496×10^{-5}	.9608	2.038	0.853	0.178	0.132
Coffee, Tea, & Cocoa	.8839	.2512	9.223×10^{-3}	.859	2.148	15.97**	1.053**	3.013**

* indicates statistically different from zero at the 5% significance level; ** at the 1% significance level.

Table 2. Demand Model Summary Statistics and Diagnostics.

	R^2	$\sqrt{T}\,\bar{\hat{\varepsilon}}_i/\hat{\sigma}_i$	$\max_{0\le z\le 1}\lvert\hat{B}_{iT}(z)\rvert$	D–W
Milk & Cream	.9856	.0134	.312	2.121
Butter	.9923	.0069	.406	1.084**
Cheese	.9946	.0128	.339	2.134
Frozen Dairy	.9670	-.0036	.543	1.526
Other Dairy	.7244	-.0192	.291	2.214
Beef	.9910	-.0007	.557	1.331*
Pork	.9668	.0057	.491	1.384*
Other Red Meat	.9598	4.9×10^{-5}	.415	1.895
Fish	.9922	-.0110	.557	1.619
Poultry	.9876	.0059	.606	1.131**
Fresh Citrus Fruit	.7286	-.0025	.574	1.818
Fresh Non-citrus Fruit	.9451	-.0149	.399	2.846
Fresh Vegetables	.9906	-.0062	.222	2.599
Potatoes	.9446	.0121	.644	2.046
Processed Fruit	.9888	-.0165	.517	1.689
Processed Vegetables	.9767	-.0153	.346	1.905
Fats & Oils	.9622	-.0133	.488	1.374
Eggs	.9977	-2.4×10^{-4}	.487	1.779
Pasta & Cereal Grains	.9843	5.3×10^{-4}	.577	1.509
Sugar & Sweeteners	.9931	-.0065	.426	2.590
Coffee, Tea, & Cocoa	.9738	.01824	.377	1.767

R^2 is the squared correlation of the observed and predicted dependent variable; $\sqrt{T}\,\bar{\hat{\varepsilon}}_i/\hat{\sigma}_i \overset{a}{\sim} n(0,1)$ is an asymptotic test statistic for zero means of the demand residuals; $\max_{0\le z\le 1}\lvert\hat{B}_{iT}(z)\rvert$ is a single equation tied Brownian bridge test; D-W is the Durbin-Watson test statistic for serial correlation in the error terms. * indicates statistically different from the null hypothesis at the 5% significance level; ** at the 1% significance level

Table 3. Constants, demographics, main diagonal of *B, Box-Cox, and rank 3 parameters.**

Parameter	Estimate	Asymptotic Standard Error	Asymptotic t–ratio	Asymptotic P–value
Dairy Products				
Milk & Cream				
Constant	400.4516	167.0079	2.397801	.016
$Age_{<5}$	–385.3338	430.3245	–.895449	.371
Age_{5-14}	861.7038	747.7986	1.152321	.249
Age_{15-24}	–268.2246	290.2010	–.924272	.355
$Age_{>54}$	1091.170	819.8913	1.330872	.183
Black	–49.00172	39.87195	–1.228977	.219
Not White or Black	–1.171513	27.22329	–.043034	.966
Own price $b_{1,1}$	2.398551	1.029731	2.329299	.020
Butter				
Constant	27.56409	8.613418	3.200134	.001
$Age_{<5}$	–5.864814	19.97749	–.293571	.769
Age_{5-14}	19.90548	36.44433	.546189	.585
Age_{15-24}	15.60141	14.13944	1.103397	.270
$Age_{>54}$	–1.758622	32.55507	–.054020	.957
Black	–2.788001	1.959402	–1.422883	.155
Not White or Black	1.300108	1.566214	.830096	.406
Own price $b_{2,2}$	$.2986517 \times 10^{-2}$	$.1283688 \times 10^{-2}$	2.326513	.020
Cheese				
Constant	–4.221629	9.813378	–.430191	.667
$Age_{<5}$	–10.63126	30.35714	–.350206	.726
Age_{5-14}	78.59921	51.25275	1.533561	.125
Age_{15-24}	–2.684196	21.86788	–.122746	.902
$Age_{>54}$	120.4936	60.81435	1.981336	.048
Black	–2.306349	2.653786	–.869079	.385
Not White or Black	1.997360	1.988838	1.004285	.315
Own price, $b_{3,3}$	$.6855164 \times 10^{-2}$	$.3665781 \times 10^{-2}$	1.870042	.061
Frozen Dairy				
Constant	16.62731	23.63151	.703607	.482
$Age_{<5}$	172.4398	94.52512	1.824275	.068
Age_{5-14}	169.7919	118.9071	1.427938	.153
Age_{15-24}	117.6385	65.39845	1.798797	.072
$Age_{>54}$	249.1209	125.5670	1.983968	.047
Black	–11.42917	6.839709	–1.671002	.095
Not White or Black	10.30628	5.645212	1.825668	.068
Own price, $b_{4,4}$	.0352694	.0231841	1.521275	.128
Canned & Powdered Milk				
Constant	77.02498	58.30862	1.320999	.187
$Age_{<5}$	172.5077	178.1571	.968290	.333
Age_{5-14}	–250.5331	238.7480	–1.049362	.294
Age_{15-24}	41.47852	117.4864	.353050	.724
$Age_{>54}$	–355.3721	276.5997	–1.284788	.199
Black	.9626987	13.54746	.071061	.943
Not White or Black	5.514632	10.28109	.536386	.592
Own price, $b_{5,5}$	.3047675	.1294329	2.354637	.019

Table 3, continued.

Parameter	Estimate	Asymptotic Standard Error	Asymptotic t–ratio	Asymptotic P–value
Meats, Fish, and Poultry				
Beef				
Constant	–98.71441	25.39640	–3.886946	.000
$Age_{<5}$	–13.25319	40.70461	–.325594	.745
Age_{5-14}	35.92908	81.73154	.439599	.660
Age_{15-24}	–54.77663	32.14665	–1.703961	.088
$Age_{>54}$	59.69787	79.72590	.748789	.454
Black	12.18492	4.146443	2.938643	.003
Not White or Black	–10.83663	3.367013	–3.218469	.001
Own price, $b_{6,6}$	.0305656	$.8244024 \times 10^{-2}$	3.707607	.000
Pork				
Constant	70.80908	21.72403	3.259482	.001
$Age_{<5}$	98.57318	69.29915	1.422430	.155
Age_{15-24}	103.0047	95.40777	1.079626	.280
Age_{15-24}	44.34759	47.89064	.926018	.354
$Age_{>54}$	115.6786	96.41663	1.199779	.230
Black	–9.921528	5.758482	–1.722942	.085
Not White or Black	3.383852	4.227421	.800453	.423
Own price, $b_{7,7}$	.0458794	.0126597	3.624049	.000
Other Red Meat				
Constant	13.81569	8.957978	1.54227	.123
$Age_{<5}$	–16.22974	24.98670	–.649535	.516
Age_{5-14}	27.38998	38.15348	.717889	.473
Age_{15-24}	–8.048191	17.77640	–.452746	.651
$Age_{>54}$	7.480517	39.55528	.189116	.850
Black	–1.009080	2.072770	–.486827	.626
Not White or Black	–.1877308	1.579160	–.118880	.905
Own price, $b_{8,8}$	.0156278	$.5855231 \times 10^{-2}$	2.669028	.008
Fish & Shellfish				
Constant	17.14271	11.91130	1.439197	.150
$Age_{<5}$	16.01230	37.58487	.426030	.670
Age_{5-14}	71.29501	47.99914	1.485340	.137
Age_{15-24}	19.72345	27.77379	.710146	.478
$Age_{>54}$	32.51503	44.13425	.736730	.461
Black	–3.476629	2.797261	–1.242869	.214
Not White or Black	3.993370	2.303270	1.733783	.083
Own price, $b_{9,9}$	$.4395113 \times 10^{-2}$	$.2102867 \times 10^{-2}$	2.090057	.037
Poultry				
Constant	9.424577	19.41054	.485539	.627
$Age_{<5}$	25.86903	59.61720	.433919	.664
Age_{5-14}	112.7465	93.17606	1.210037	.226
Age_{15-24}	–1.192216	43.57236	–.027362	.978
$Age_{>54}$	34.94146	80.64696	.4332644	.665
Black	–2.948940	4.892394	–.602760	.547
Not White or Black	10.53016	4.991436	2.109646	.035
Own price, $b_{10,10}$	.0181704	$.5140193 \times 10^{-2}$	3.534961	.000

Table 3, continued.

Parameter	Estimate	Asymptotic Standard Error	Asymptotic t–ratio	Asymptotic P–value
Fruits and Vegetables				
Fresh Citrus Fruit				
Constant	151.0950	42.01749	3.596002	.000
$Age_{<5}$	–180.4505	87.20089	–2.069366	.039
Age_{5-14}	–222.5962	140.6704	–1.582395	.114
Age_{15-24}	–106.4208	61.22153	–1.738291	.082
$Age_{>54}$	–115.9288	123.5834	–.938061	.348
Black	–4.129310	7.170613	–.575866	.565
Not White or Black	–6.175232	5.902981	–1.046121	.296
Own price, $b_{11,11}$	.0563934	.0200630	2.810819	.005
Fresh Non-citrus Fruit				
Constant	153.7503	73.62322	2.088340	.037
$Age_{<5}$	138.1371	193.4211	.714178	.475
Age_{5-14}	–136.8518	262.7964	–.520752	.603
Age_{15-24}	95.30952	126.8364	.751437	.452
$Age_{>54}$	–438.7356	286.0211	–1.533927	.125
Black	–5.940248	15.88149	–.374036	.708
Not White or Black	10.75254	12.71678	.845540	.398
Own price, $b_{12,12}$	.1718975	.1015319	1.693039	.090
Fresh Vegetables				
Constant	225.0974	61.53353	3.658126	.000
$Age_{<5}$	–146.0028	117.4012	–1.243622	.214
Age_{5-14}	122.4797	160.5392	.762927	.446
Age_{15-24}	–17.89469	82.24098	–.217589	.828
$Age_{>54}$	178.0454	175.4719	1.014666	.310
Black	–22.40083	10.86436	–2.061862	.039
Not White or Black	14.74307	8.446066	1.745554	.081
Own price, $b_{13,13}$	.1162223	.0571378	2.034071	.042
Potatoes				
Constant	159.8164	68.42335	2.335700	.020
$Age_{<5}$	–71.66518	176.5486	–.405923	.685
Age_{5-14}	–104.1668	242.9075	–.428833	.668
Age_{15-24}	41.54775	132.3913	.313825	.754
$Age_{>54}$	–428.2641	250.0232	–1.712897	.087
Black	–2.260539	15.15755	–.149136	.881
Not White or Black	–1.981753	11.94165	–.165953	.868
Own price, $b_{14,14}$	.4476947	.1747796	2.561481	.010
Processed Fruit				
Constant	–31.29244	25.11612	–1.245911	.213
$Age_{<5}$	–98.26029	83.44195	–1.177589	.239
Age_{5-14}	199.1655	128.8688	1.545490	.122
Age_{15-24}	–70.18885	62.04413	–1.131273	.258
$Age_{>54}$	215.4645	121.9345	1.767051	.077
Black	.0759863	6.138067	.012380	.990
Not White or Black	4.631180	5.387631	.859595	.390
Own price, $b_{15,15}$	.0231164	$.7450056 \times 10^{-2}$	3.102843	.002

Table 3, continued.

Parameter	Estimate	Asymptotic Standard Error	Asymptotic t–ratio	Asymptotic P–value
Processed Vegetables				
Constant	51.82672	33.75945	1.535177	.125
$Age_{<5}$	34.91193	101.2193	.344914	.730
Age_{5-14}	65.53713	148.6315	.440937	.659
Age_{15-24}	80.31124	77.52784	1.035902	.300
$Age_{>54}$	79.09137	158.9078	.497719	.619
Black	–6.974044	8.805015	–.792054	.428
Not White or Black	7.289670	7.169072	1.016822	.309
Own price, $b_{16,16}$	.1067930	.0504515	2.116744	.034
Miscellaneous Foods				
Margarine, Fats & Cooking Oils, Excluding Butter				
Constant	8.809842	15.57483	.565646	.572
$Age_{<5}$	–7.119928	45.23274	–.157407	.875
Age_{5-14}	4.398767	79.31484	.055460	.956
Age_{15-24}	17.53647	33.01583	.531153	.595
$Age_{>54}$	–6.224789	78.23397	–.079566	.937
Black	.2162304	4.254941	.050819	.959
Not White or Black	3.446176	3.602691	.956556	.339
Own price, $b_{17,17}$	.0174848	$.7435984 \times 10^{-2}$	2.351374	.019
Eggs				
Constant	58.21888	17.25044	3.374922	.001
$Age_{<5}$	148.8485	56.92159	2.614974	.009
Age_{5-14}	58.48786	84.66879	.690784	.490
Age_{15-24}	73.39627	34.02838	2.156914	.031
$Age_{>54}$	136.1283	85.93384	1.584106	.113
Black	–10.10603	4.827193	–2.093562	.036
Not White or Black	4.587329	3.698107	1.240453	.215
Own price, $b_{18,18}$	.0165967	$.6470346 \times 10^{-2}$	2.565036	.010
Bread, Cereal Grain & Pasta				
Constant	260.0493	98.41350	2.642414	.008
$Age_{<5}$	–347.3041	180.6362	–1.922672	.055
Age_{5-14}	–274.3420	276.1288	–.993529	.320
Age_{15-24}	–194.5826	129.3603	–1.504191	.133
$Age_{>54}$	–1153.038	375.7599	–3.068550	.002
Black	10.90580	17.09694	.637880	.524
Not White or Black	4.202203	11.59157	.362522	.717
Own price, $b_{19,19}$	.3584566	.2843467	1.260632	.207
Sugar and Caloric Sweeteners				
Constant	118.5841	33.34448	3.556334	.000
$Age_{<5}$	–53.61066	67.72584	–.791584	.429
Age_{5-14}	16.68190	100.3731	.166199	.868
Age_{15-24}	56.35774	58.10219	.969976	.332
$Age_{>54}$	–71.84597	94.95358	–.756643	.449
Black	–6.877342	6.466937	–1.063462	.288
Not White or Black	5.834398	5.500398	1.060723	.289
Own price, $b_{20,20}$	.0782676	.0242628	3.225819	.001

Table 3, continued.

Parameter	Estimate	Asymptotic Standard Error	Asymptotic t–ratio	Asymptotic P–value
Coffee, Tea, & Cocoa				
Constant	36.71345	9.086131	4.040603	.000
$Age_{<5}$	–8.313453	18.27687	–.454862	.649
$Age_{5\text{-}14}$	29.30088	29.54195	.991840	.321
$Age_{15\text{-}24}$	–2.610791	14.29012	–.182699	.855
$Age_{>54}$	59.27300	32.86510	1.803524	.071
Black	–4.343055	1.946413	–2.231312	.026
Not White or Black	1.183474	1.318227	.897778	.369
Own price, $b_{21,21}$	$.2820930\text{x}10^{-2}$	$.7931703\text{x}10^{-3}$	3.556525	.000
Nonfood Expenditure				
Constant	520.8642	1205.295	.432147	.666
$Age_{<5}$	4825.422	5299.583	.910529	.363
$Age_{5\text{-}14}$	16187.01	6824.572	2.371872	.018
$Age_{15\text{-}24}$	5094.378	3966.541	1.284338	.199
$Age_{>54}$	16626.07	6503.471	2.556491	.011
Black	–667.722	355.6175	–1.877641	.060
Not White or Black	752.4342	323.6012	2.325190	.020
Own price, $b_{22,22}$	1.000000	——	——	——
Box-Cox Parameters				
κ	.8584773	.0285454	30.07414	.000
λ	1.088438	.0332934	32.69235	.000
Quadratic Rank 3 Parameters				
δ_{Milk}	$-.7979284\times10^{-05}$	$.1231113\times10^{-04}$	–.648136	.517
δ_{Butter}	$.6516140\times10^{-06}$	$.6331848\times10^{-06}$	1.029105	.303
δ_{Cheese}	$-.8400016\times10^{-07}$	$.7822616\times10^{-06}$	–.107381	.914
$\delta_{Ice\ Cream}$	$.2378841\times10^{-05}$	$.1707076\times10^{-05}$	1.393518	.163
$\delta_{Other\ Dairy}$	$.5046246\times10^{-05}$	$.4725501\times10^{-05}$	1.067875	.286
δ_{Beef}	$-.4663525\times10^{-05}$	$.1707923\times10^{-05}$	–2.730524	.006
δ_{Pork}	$-.4274155\times10^{-05}$	$.1552887\times10^{-05}$	–2.752392	.006
$\delta_{Other\ Meat}$	$.3891047\times10^{-06}$	$.7430648\times10^{-06}$	.523648	.601
δ_{Fish}	$.8892863\times10^{-06}$	$.5139223\times10^{-06}$	1.730390	.084
$\delta_{Poultry}$	$.1273841\times10^{-05}$	$.1347266\times10^{-05}$	.945501	.344
$\delta_{Fresh\ Citrus}$	$-.6101568\times10^{-05}$	$.2460191\times10^{-05}$	–2.480120	.013
$\delta_{Fresh\ Noncitrus}$	$-.8740211\times10^{-06}$	$.4697163\times10^{-05}$	–.186074	.852
$\delta_{Fresh\ Vegetables}$	$-.2427044\times10^{-05}$	$.2756051\times10^{-05}$	–.880624	.379
$\delta_{Potatoes}$	$-.9658178\times10^{-05}$	$.4938488\times10^{-05}$	–1.955696	.051
$\delta_{Processed\ Fruit}$	$.5092494\times10^{-05}$	$.2000672\times10^{-05}$	2.545392	.011
$\delta_{Processed\ Vegetables}$	$.2361664\times10^{-05}$	$.2700797\times10^{-05}$	.874432	.382
$\delta_{Fats\ \&\ Oils}$	$.8415340\times10^{-06}$	$.1300952\times10^{-05}$	.646860	.518
δ_{Eggs}	$-.8110274\times10^{-06}$	$.1335823\times10^{-05}$	–.607137	.544
δ_{Cereas}	$.5792619\times10^{-05}$	$.5098581\times10^{-05}$	1.136124	.256
δ_{Sugar}	$-.1684004\times10^{-05}$	$.1846147\times10^{-05}$	–.912172	.362
δ_{Coffee}	$-.1364496\times10^{-05}$	$.4766596\times10^{-06}$	–2.862622	.004

Figures

Figure 1a. US per Capita Real Income Distribution Parameters.

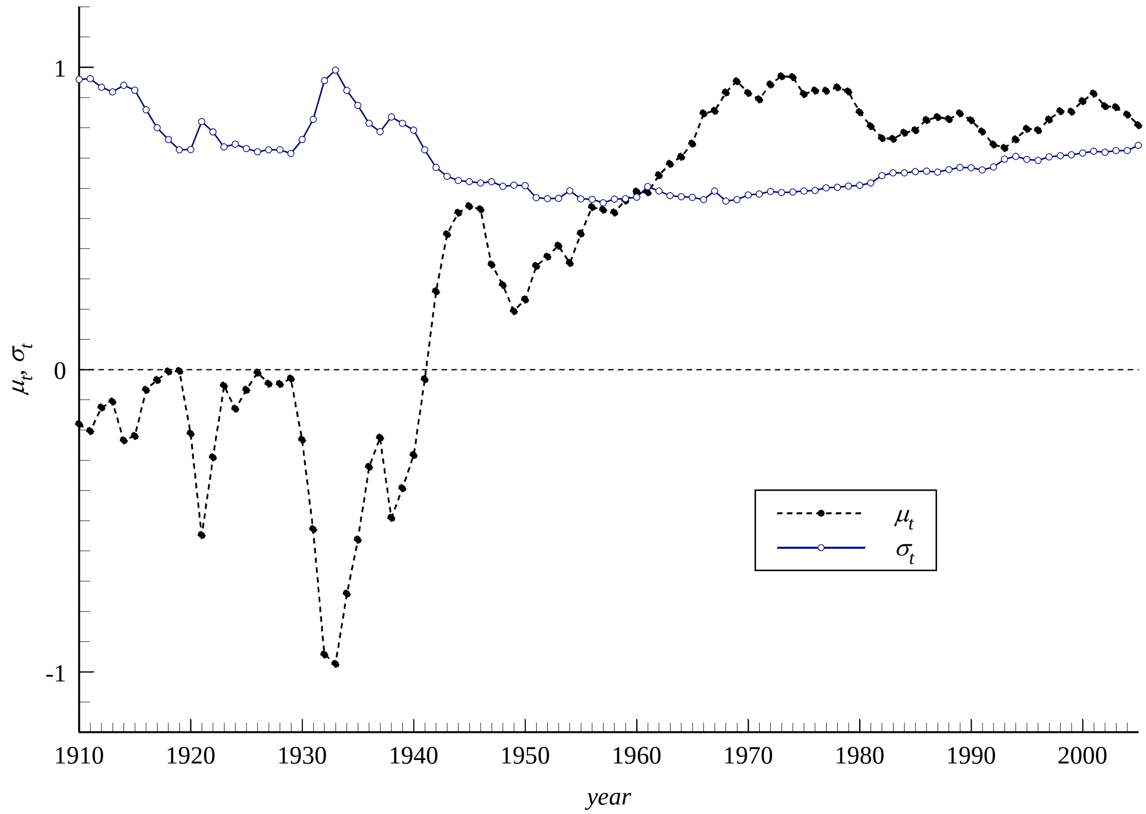

Figure 1b. Distribution of U.S. Real Family Income, 1910-2000.

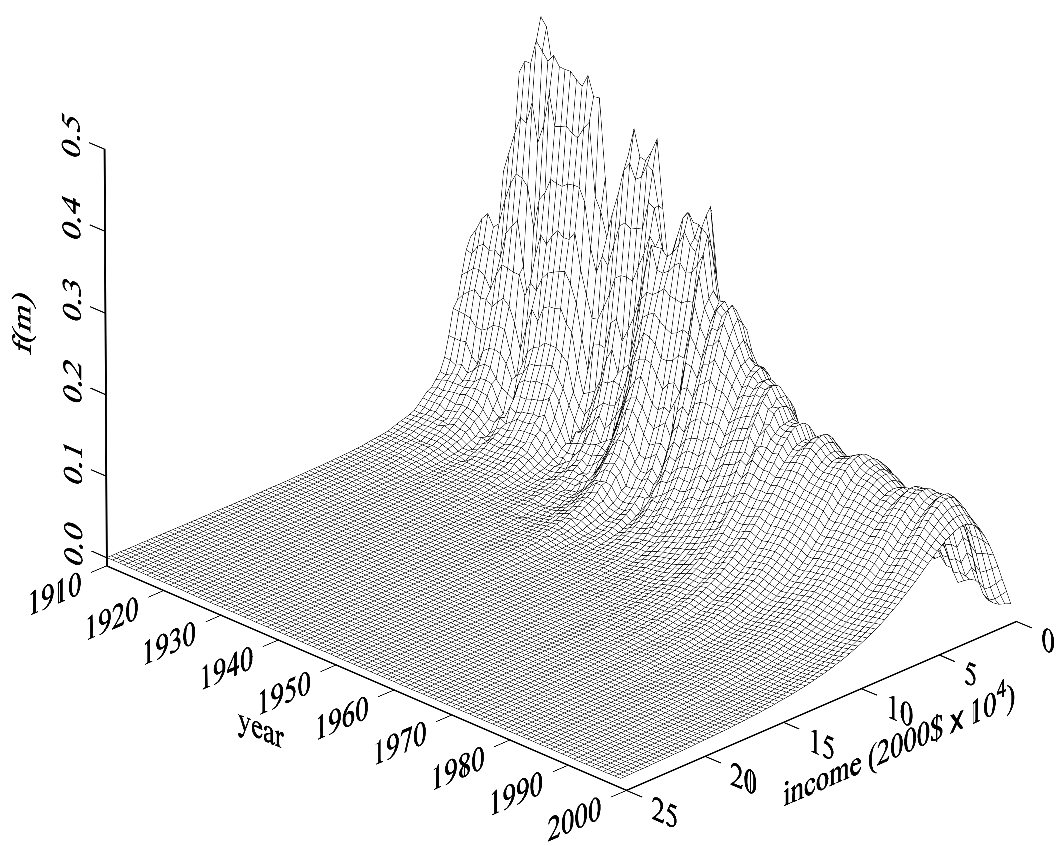

Figure 2a. US Population Percentages by Race.

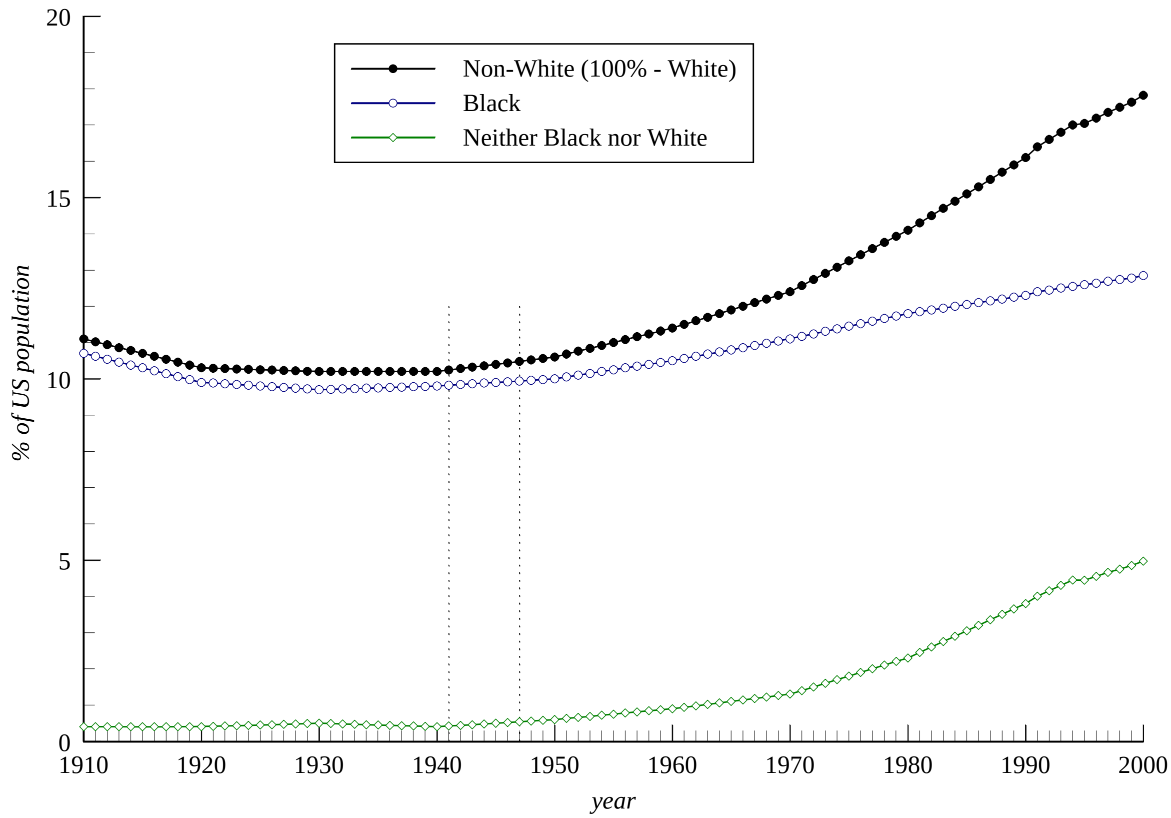

Figure 2b. US Population Proportions by Age Group.

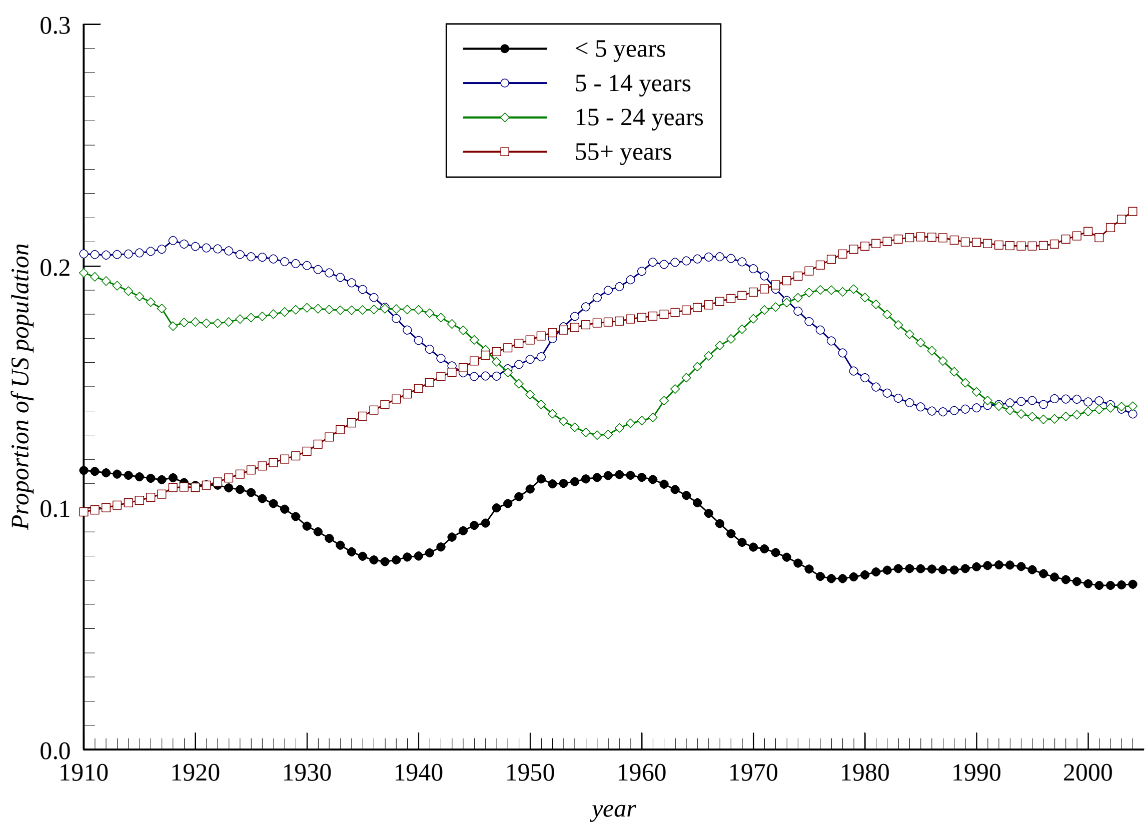

Figure 3. Conditional NL3SLS Search over κ and λ.