A Model of Food Demand, Nutrition and the Effects of Agricultural Policy

Timothy K.M. Beatty

and

Jeffrey T. Lafrance

DRAFT

1.0 Introduction

Farm and food policy in the United States is in the midst of a major transformation. The 1996 Food and Agriculture Improvement Act of 1996 replaced most, though not all, farm-level price and income support programs with direct cash payments. At the same time, welfare, food stamps, Women, Infants and Children, Aid to Families with Dependent Children, and school lunch programs have been reduced in scope at the Federal level through block grants to states and increased flexibility by states to design individual policies. With the passage of HR 2559, subsidized crop insurance has become increasingly important as a mechanism for supporting the incomes of U.S. farmers. These policy changes all influence the prices paid for and quantities consumed of food items and therefore implicitly affects the prices paid for and quantities consumed of nutrients. These policies also affect the incomes and food expenditures of U.S. consumers. Exactly how much and in what direction these effects are realized, however, is an important and relevant policy question.

By and large, past farm-level policies (i.e., price supports and other commodity programs) created consumer incentives that oppose those created by food subsidy programs. While food aid recipients spend more on food, they may eat less healthy foods due to relative price distortions created by farm-level policies. On the other hand, people that are neither farmers nor food aid recipients pay higher taxes to finance farm and food subsidies. This lowers disposable incomes, food expenditures, and economic welfare. In addition, policy-induced price distortions can, and often do, create incentives to consume a less healthy mix of foods for members of this group. This paper will develop improved methods for addressing these issues on the consumer side of the food sector and increase the level of understanding of the consumption, economic welfare, and nutritional impact of past and present U.S. farm and food policies.

In this paper we propose and implement a methodology by which the impacts of agricultural policies on food consumption, nutrition and consumer welfare can be coherently measured. A method for nesting and testing the functional form of the income terms in an incomplete system of aggregable and integrable demand equations will be derived by nesting the Gorman Polar Form (Gorman) demand model within a Price Independent Generalized Linear Incomplete Demand System (LaFrance and Hanemann; Lewbell 1987, 1990; Muellbauer 1975). When combined with information on the distribution of income, this class of demand models aggregates across income, demographic variables, and variations in micro demand parameters.

In order to infer the year-to-year U.S. income distribution using annual time series data on the ranges for quintiles and the top five percentile of the income distribution plus mean income within each range we develop a technique utilizing a truncated three-parameter lognormal distribution. Estimates of the year-to-year income distribution will be combined with annual time series data on the U.S. consumption of and retail prices for twenty-one food items to estimate the rank and functional form of the income terms in

U.S. food demand. The model will be estimated with per capita U.S. consumption of 21 food items over the period 1919-1995, using the mean, variance, and skewness of the U.S. population's age distribution and the proportion of the population that is White, Black, and neither White nor Black as demographic variables.

Similar techniques will be used to infer the share of 17 nutrients contributed by the 21 food items over the period 1909-1998 based on an extraneous sample of data for the nutrient content shares for the years 1952-1982. These nutrient content share estimates will be combined with the food demand model to estimate the price and income response elasticities of U.S. nutrient consumption across ethnic, age, and income groups of the population. These elasticities will then be used to infer the nutritional impacts of Federal farm and food aid policies during the last half of the 20th century, as well as distributional effects on consumer economic welfare of these policies.

2.0 Data

In order to answer the questions posed above, we will employ four different time series data sets. The first is data on per capita consumption of food items and their corresponding prices. Currently, this data set consists of annual time series observations over the period 1919-1995. Per capita consumption of twenty-one food items and corresponding average retail prices for those items were constructed from several USDA and Bureau of Labor Statistics sources. The quantity data are aggregates taken from the USDA series *Food Consumption, Prices and Expenditures*. Estimated retail prices corresponding to the quantity data were constructed for 1967 using detailed disaggregated retail price estimates along with the respective quantity observations to construct an average retail price per pound in 1967 for each food category. For all other years, the fixed 1967 consumer price indices or average retail food prices were combined to construct a consistent retail price series for each commodity. The consumer price index for nonfood items is used as the price of nonfood expenditures.

The second data series are demographic factors that explain demand, but will also be used to distinguish the nutritional impact of different policies on different demographic groups in the population. These demographic factors include the first three central moments (i.e., the mean, variance, and skewness) of the age distribution and the proportion of the U.S. population that is Black and the proportion that is neither Black nor White. The estimated age distribution is based on ten-year age intervals, plus categories for children less than five years old and adults that are sixty-five years old and older. The ethnic variables are linearly interpolated estimates of Bureau of Census figures reported on ten-year intervals. Habit formation will also be entertained and empirically investigated by including lagged quantities demanded among the demographic variables in the empirical estimation process.

The third data series involves the U.S. income distribution. The Bureau of the Census publishes annually quintile ranges, intra-quintile means, the top five-percentile lower bound for income, and the mean income within the top five-percentile range for all U.S. families. These data are available for the years 1929, 1935, 1936, 1941 and 1946 from the Census Bureau's historical statistics. The data for the years 1947-1998 are available over the Internet from the Census Bureau. In order to estimate the income limits for the missing years, we take the following approach: To estimate the lowest income limit (m_{1t}

the upper limit of the lowest quintile, or equivalently the lower limit of the second quintile), we utilize data on per capita disposable income and the unemployment rate as predictors. For the other income limits, we follow a recursive forecasting procedure in which each limit is a function of a third order polynomial of the income limit below it.

Finally, we currently have annual estimates of the percentages of the total availability of seventeen nutrients from each of the 21 food items and 17 nutrients for the period 1952-1982. These nutrient content data will be augmented with USDA Center for Nutrition Policy and Promotion nutrient content data for the 21 foods and 17 nutrients for the period 1909-1998 as soon as feasible.

3.0 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. We propose the development and implementation of a method by which the impact of a given policy on nutrient intake can be evaluated. The purposes of this new approach are to derive and implement a theoretically consistent empirical model of household food consumption that: (1) nests the functional form of the income terms in the food demand equations; (2) when combined with information on the income distribution, permits consistent and asymptotically efficient estimation of the parameters of the demand equations with aggregate data; (3) when combined with inferences and/or data on the nutritional content of individual food items, permits inferences on the nutritional impacts of farm and food policies on consumers as a whole and on income and demographic groups of consumers of particular interest to policy-makers and analysts; and (4) permits inferences on the economic welfare impacts of farm and food policies on consumers as a whole and on income and demographic groups of consumers of particular interest to policy-makers and analysts.

3.1 A Nested Incomplete Demand System for Food

We start with a stable, theoretically consistent reduced form econometric model of demands for food items with conditional mean given by, $E(q\,|\,p,m,d) = h(p,m,d)$, where q is an n_q-vector of food quantities, p is an n_q-vector of food prices, m is income and d is a k-vector demographic characteristics. Let x denote the scalar variable for total consumer expenditures on all nonfood items. Assume that each of the prices for individual food items and income are deflated by a price index measuring the cost of nonfood items. Consider the Gorman Polar Form for the (quasi-) indirect utility function generated by a quadratic (quasi-)utility function,

$$v(p,m,d) = \frac{(m - \alpha(d)'p - \alpha_0(d))}{\sqrt{p'Bp + 2\gamma'p + \gamma_0}}, \qquad (1.1)$$

where $\alpha(d)$ is an n_q-vector of functions of the demographic variables, $\alpha_0(d)$ is a scalar function of the demographic variables, B is an $n_q \times n_q$ matrix of parameters, γ is an n_q-vector of parameters, and γ_0 is a scalar parameter. For identification purposes, we choose the normalization $\gamma_0 = 1$.

Applying Roy's identity to this (quasi-) indirect utility function generates a PIGL-IDS system of demands.

$$E(q \mid p, m, d) = \alpha + \frac{(m - \alpha(d)' p - \alpha_0(d))}{(p' B p + 2\gamma' p + 1)} (Bp + \gamma) .$$

Next, define Box-Cox transformations on m and p by $m(\kappa) = (m^{\kappa} - 1)/\kappa$ and $p_i(\lambda) = (p_i^{\lambda} - 1)/\lambda$, for $i = 1, \ldots, n_q$, with $p(\lambda) \equiv [p_1(\lambda), \ldots, p_n(\lambda)]'$, and replace m and p with $m(\kappa)$ and $p(\lambda)$, respectively, in (1.1). Applying Roy's identity to the resulting (quasi-) indirect utility function then gives a PIGL-IDS that can be written in expenditure form as,

$$E(e \mid p, m, d) = P^{\lambda} m^{1-\kappa} \left\{ \alpha(d) + \left(\frac{m(\kappa) - \alpha(d)' p(\lambda) - \alpha_0(d)}{p(\lambda)' B p(\lambda) + 2\gamma' p(\lambda) + 1} \right) (Bp(\lambda) + \gamma) \right\}, \quad (1.2)$$

where $e = [\, p_1 q_1 \cdots p_n q_n \,]'$ is the n_q-vector of (deflated) expenditures on the food items q. This is the form in which the empirical demand analysis will be conducted. The fundamental empirical issue regarding the income term to be addressed in this paper, concerns the point estimate for $\kappa \in [0, \infty)$. Additional hypotheses of interest include the existence of habit formation (i.e., whether or not lagged quantities demanded, q_{-1}, are components of the "demographic effects"), and whether or not demand for food is separable from expenditure on all other goods. We will consider each of these questions in turn in a subsequent section.

4.0 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 N denote an $n_z \times n_q$ 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) N, we can analyze policy effects on nutritional intakes using the previously described demand model since

$$z = Nq \ , \tag{1.3}$$

where z is an n_z-vector of nutrients important to the household.

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. These elasticities satisfy

$$\boldsymbol{e}_{p_k}^{z_i} = \sum_{j=1}^{n_q} s_{ij} \boldsymbol{e}_{p_k}^{q_j} \ , \tag{1.4}$$

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

$$\boldsymbol{e}_{m}^{z_i} = \sum_{j=1}^{n_q} s_{ij} \boldsymbol{e}_{m}^{q_j} \ , \tag{1.5}$$

where $\boldsymbol{e}_{m}^{q_j} \equiv (m / q_j) \cdot (\partial q_j / \partial m)$ is the income elasticity of food j. One goal of this paper will be to obtain accurate estimates and/or measures of the matrices $[s_{ij}]$, $[\boldsymbol{e}_{p_k}^{z_i}]$ and $[\boldsymbol{e}_{m}^{z_i}]$.

5.0 Inferring the U.S. Income Distribution

5.1 Theory

The demand model described above is nonlinear in income. Therefore, the demand equations do not aggregate directly across individuals to average income at the market level. The advantage of using the Gorman class of Engel curves is that to generate a theoretically consistent, aggregable model of demand, only a limited number of statistics concerning the income distribution are needed. The demand model proposed in this paper requires two moments of the income distribution, specifically those associated with m^{1-k} and m .

The U.S. Bureau of the Census publishes annually quintile ranges, intra quintile means, the top five-percentile lower bound for income, and the mean income within the top five-percentile range for all U.S. families. To integrate this data efficiently with the demand model developed in the previous section, a reasonable estimate of the entire income distribution function must be generated from these summary statistics. The simplest approach would be to construct a sequence of piecewise uniform densities on each of the five ranges. However, such a piecewise uniform distribution would contradict some of the available information. In particular, it generally will not satisfy the mean conditions in each of the intervals. In addition, this density function is discontinuous at six (6) points arbitrarily determined by the way in which the Census Bureau reports the income distribution data.

One solution is to construct a pair of uniform densities over each of the ranges, separated at the intra-range mean, with probability weights for each sub-range that maintain the within quintile conditional means and have total probability that sums to the appropriate quintile or top-five percent weight (i.e., .20 or .05, respectively). However this approach increases the number of discontinuities from six (6) to twelve (12), which again are pure artifacts of the way in which the data is reported.

An alternative approach is to fit a continuous density function to the data. Given the shape of the above piecewise uniform income distribution – namely, unimodal and right skewed - an obvious choice is the lognormal distribution. However the lognormal has been criticized for estimating income distributions (McDonald). Therefore the proposed approach for this paper is to estimate the income distribution using a truncated three-parameter lognormal distribution. Let the random variable z have a standard normal distribution. Define the latent income variable m by $\ln(m-\theta)=\mu+\sigma z$ where $m>\theta$ and $\{\mu,\sigma,\theta\}$ are constants. Then m has a three-parameter lognormal distribution with density function

$$f(m;\mu,\sigma,\theta)=\frac{1}{\sqrt{2\pi}\sigma(m-\theta)}\exp\left\{-\frac{1}{2\sigma^2}[\ln(m-\theta)-\mu]^2\right\}. \tag{1.6}$$

As previously noted, we only observe values of $m>0$. Therefore, define the standardized zero income limit, $z_0=(\ln(-\theta)-\mu)/\sigma$. The truncated three-parameter lognormal can then be written as

$$f(m\mid m\geq 0;\mu,\sigma,\theta)=\frac{1}{\sqrt{2\pi}\sigma(m-\theta)(1-\Phi(z_0))}\exp\left\{-\frac{1}{2\sigma^2}[\ln(m-\theta)-\mu]^2\right\}. \tag{1.7}$$

where Φ is the standard normal distribution function.

Since the data takes the form of income limits m_i, where $\Pr(m \le m_i \mid m \ge 0) = F_i = .2,.4,.6,.8,.95$ for $i = 1,2,3,4,5$, respectively, to estimate the truncated three-parameter lognormal we define standardized limits z_i (equivalent to z_0) by

$$z_i \equiv \frac{\ln(m_i - \theta) - \mu}{\sigma},\ i = 1 \ldots 5. \tag{1.8}$$

This allows us to write an empirical limit equation that we can use to estimate $\{\mu, \sigma, \theta\}$ in the form,

$$F_i = \frac{\Phi(z_i)}{1 - \Phi(z_0)} + \varepsilon_i,\ i = 1 \ldots 5, \tag{1.9}$$

where ε_i is an estimation error. Given estimates for $\{\mu, \sigma, \theta\}$ we can numerically integrate equation (1.7) using Simpson's rule to obtain estimates of the required moments (m^{1-k}, m) of the income distribution to an arbitrary level of precision.

5.2 Estimation Results

Figure 1 shows estimates of the mean (μ,mu), the variance(σ ,sigma) and the truncation parameter (θ,theta) for each year over the period 1909-1999. We note that save for a period during the depression and the war years, the variance parameter has remained approximately stable. This would seem to suggest that the dispersion of real income has remained roughly constant over the last century. The mean parameter is steadily increasing throughout the period, again with the notable exception of the depression and the war years. This is consistent with mean household income increasing in real terms over time. Perhaps the most interesting feature of Figure 1 is the behaviour of the truncation parameter, which declines steadily over the entire period. This suggests that the share of the population with a negative household income has been increasing over time. The fact that theta declines over time may also reflect a problem in the way the Bureau of the Census collects and reports household income. The Bureau of the Census does not report any households with negative income, implicitly restricting the distribution of income to the positive orthant.

6.0 Inferring Nutrient Shares

6.1 Theory

We currently have information obtained from the USDA's Human Nutrition Information Service on the nutrient content of 21 food items over the period 1952-1982. The USDA has updated the nutrient availability data several times since the sample of 31 annual observations was initially constructed and these nutrient content estimates are no longer consistent with the available data on annual food and nutrient consumption. In order to work with the data we currently have, we assume that nutrient content of foods varies smoothly over time. Since the quantity of the i^{th} nutrient consumed is a linear function of the nutrient content of foods and the quantity of the individual foods consumed, $z_{it} = \sum_{j=1}^{n_q} n_{ijt} q_{jt}$, $i = 1,\ldots,n_z$, $t = 1,\ldots,T$, the share of nutrient i from food item j is $s_{ijt} = n_{ijt} q_{jt} / z_{it}$, $i = 1,\ldots,n_z$, $j = 1,\ldots,n_q$, and $t = 1,\ldots,T$. Each of these shares is non-negative and the shares of nutrient i contributed by all foods always sum to one. A simple specification that maintains this relationship in conditional means and in levels of the nutrient shares, both in the sample period and for out of sample forecasts is the multinomial logit,

$$s_{ijt} = \frac{e^{b_{ij}(t)}}{\sum_{j=1}^{n_q} e^{b_{ij}(t)}} + e_{ijt}, \ i = 1,...,n_z, \ \ j = 1,...,n_q, \ t = 1,...,T, \qquad (1.10)$$

where $b_{ij}(t)$ is assumed to be a smooth function of t that can be reasonably approximated by a low order polynomial in t and e_{ijt} is a mean zero, constant variance error term. We normalize for identification by setting $b_{i1}(t) = 0 \ \forall i, \forall t$, and estimate the remaining n_q-1 polynomial terms with nonlinear seemingly unrelated regressions equations methods.

6.2 Results

Figure 2 shows the reported and estimated percent contributions to cholesterol intake for eight food groups (Fresh Milk and Cream, Butter, Cheese, Beef and Veal, Pork, Fish, Poultry and Eggs) over the period 1952 to 1982. In terms of our previous discussion this represents s_{ijt}, where i is cholesterol, j are the foods listed above and t corresponds to 1952-1982. These percentages were calculated for all 21 food categories and 17 nutrients in each year. Figure 2 illustrates the fact that the data on nutrient shares is somewhat noisy. For example, it would seem unlikely that the share of total cholesterol consumption from pork should vary so much from year to year. A large variability in year-to-year reported nutrient shares is common throughout this data. In conjunction with the Center for Nutrition Policy and Promotion of the USDA, it is our intention to obtain better estimates of these shares in future research.

7.0 Empirical Results

We now turn our attention to considering the results of estimating the system of equations (1.2) using a two-step seemingly unrelated regression method. In order to reduce the dimensionality of the problem and to facilitate testing of the hypotheses of interest we impose symmetry on the Slutsky substitution matix. In a subsequent draft we hope to loosen this restriction and test for symmetry. In addition, we omit the data for the years 1941-1946, which reflect the impacts of U.S. participation in the Second World War. Table 1 presents some key per equation, goodness of fit results.

7.1 Demographic Translating

Our approach incorporated five demographic characteristics of the U.S. population: the mean, variance and skewness of the age distribution as well as the share of the population that is black and the share of the population that is neither black nor white. In order to infer the effect of these demographic translators on nutrient intake we proceed in the following manner: the effect of demographic translator d_{kt} on intake of nutrient i at time t, z_{it} can be written in elasticity form as:

$$\boldsymbol{e}_{z_{it}}^{d_{kt}} = \frac{d_{kt}}{z_{it}} \frac{\partial z_{it}}{\partial d_{kt}} = \sum_{j=1}^{n_j} s_{ijt} \left(\frac{d_{kt}}{q_{jt}} \frac{\partial q_{jt}}{\partial d_{kt}} \right),$$

where s_{ijt} is as before the share of nutrient i contributed by food j at time t. Figure 4 plots the elasticity of fat intake with respect to the five demographic translators over the period 1952-1982. We note that the influence of the demographic shifters appears to be constant over the period. The most striking aspect of this figure is influence of the percent of the population that is black has on fat intake. This is principally driven by significant and positive coefficients on this parameter associated with the demand for red meat, other read meat and pork. Equally worthy of note is the impact that the mean of the age distribution has on fat intake. This is driven largely by a negative coefficient on the coefficient associated with Fats and Oils and the highly negative coefficients on the coefficient associated with red meat and pork.

7.2 Functional Form

The functional form specified in this paper is extremely flexible, and allows prices and income to enter, independently, either linearly ($\boldsymbol{k} = 1, \boldsymbol{l} = 1$) or in logarithmic form ($\boldsymbol{k} = 0, \boldsymbol{l} = 0$). Figure 3 shows the sum of squared residuals for $\boldsymbol{k} \in (0,1.3)$ for both the first and second rounds of our estimation problem. We see that, except for a slight discontinuity in the neighborhood of $\boldsymbol{k} = 0.07$ and $\boldsymbol{k} = 0.08$, the sum of squared residuals is monotonically decreasing in $\boldsymbol{k}$ until it reaches a minimum at $\boldsymbol{k} = 0.92$ in the first round of estimation and $\boldsymbol{k} = 0.95$ in the second round.

We also note that the Box-Cox parameter on prices $\boldsymbol{l}$ equals .904414 and is significantly different from zero, with a p-value equal to 0.000. Thus the data strongly rejects a specification that is in terms of log-prices.

Both of these facts are extremely important in the context of applied demand analysis. They suggest that models that incorporate income and prices in logarithmic form are imposing strong restrictions upon the functional form that are not supported by the data. In a subsequent paper, we hope to explicitly nest some of the more commonly used demand systems.

7.3 Price and Income elasticities of selected nutrients

Figure 5 shows the evolution of the income elasticity of protein, carbohydrates and cholesterol, over the period 1952-1982. A significant result is that these three nutrients are all normal goods. The income elasticities of these nutrients are largely constant and positive over the period except for a short period, 1975-1979. This short period 1975-1979 is associated with a fairly general decline in the income elasticities of most major foods over this period. The low income elasticity of cholesterol is accounted for by the negative income elasticities associated with milk and butter and the very low income elasticities associated with eggs. The most significant factors in the income elasticity of protein are the slightly positive income elasticities of beef, pork and chicken over the entire period. The higher income elasticity of carbohydrates is largely due to the large share of this nutrient contributed by sugar consumption combined with positive income elasticity.

Figure 6 shows the income elasticities of vitamins A, B6 and C. The income elasticities associated with these nutrients are all slightly increasing over the period. The slight upward trend in the income elasticity of vitamin A is associated with an increase in the income elasticities of fresh vegetables and processed vegetables. The increase in the income elasticity of vitamin B6 is attributable to a decreasing share of that nutrient being obtained through consumption of potatoes (which have a negative income elasticity), and increasing income elasticities of fresh non-citrus fruit and processed vegetables. The upward trend in the income elasticity of Vitamin C is associated with increases in the income elasticities of fresh citrus fruit, fresh non-citrus fruit, fresh and processed vegetables.

Figure 7 shows the evolution of the elasticity of protein carbohydrates and cholesterol, over the period 1952-1982 with respect to the price of beef and veal. These elasticities are clearly increasing over the first half of the period and sharply decreasing over the second. This pattern is largely driven by the evolution of cross-price elasticities for foods that contribute a large percentage of the given nutrients. The gradual increase and steep decline in the elasticities associated with cholesterol are largely accounted for by an increase and subsequent decrease of the cross-price elasticity between beef and milk and the declining share of cholesterol contributed by milk consumption. In addition, the long run decrease in the cross-price elasticity between beef and eggs contributes to the sharp decline at the end of the period. A similar pattern in the price elasticity of protein can be explained by an increase and decrease in the cross-price elasticites of beef and milk, beef and cheese and beef and poultry. This pattern for carbohydrates is largely accounted for by an increase and sharp decline to almost zero in the cross-price elasticity between beef and sugar.

Figure 8 traces the evolution of the elasticity vitamin A, B6 and C with respect to the price of fresh vegetables. The fresh vegetable price elasticities of these nutrients appear to be increasing over the first half of the period and decreasing over the second. The increase and subsequent decrease in this elasticity for vitamin A can be attributed to the increase and subsequent decrease in the cross-price of elasticity fresh noncitrus fruit which can reasonably be considered substitutes. In addition the cross-price elasticity of other red meat and fresh vegetables exhibits a similar pattern. In the case of vitamin B6, this pattern appears to be driven by the cross-price elasticity of fresh non-citrus fruit and fresh vegetables and the cross-price elasticity of flour and cereals and fresh vegetables which are both increasing at the beginning of the period and decreasing at the end of the period. The price elasticity of vitamin C with respect to the price of fresh vegetables is driven by the cross price elasticities of fresh citrus fruit and fresh vegetables, fresh non-citrus fruit and fresh vegetables and potatoes and fresh vegetables.

7. 4 Separability and Habit Formation

In this section we will describe an approximate F-test proposed by Lafrance and Beatty for use with systems of equations estimated by SURE. We will use it to test several hypotheses concerning the nature of the demand for food. The test constructed by Lafrance and Beatty is an approximate F-test based on the Lagrange multiplier principle, which partially overcomes the tendency of the degrees of freedom correction to overcorrect the LM test. Let “^” denote unrestricted estimates and let “~” denote restricted estimates. Then we can write the approximate F-statistic as:

$$F(G, NT-K) \approx \frac{(\tilde{s}(\tilde{\Sigma}) - \hat{s}(\tilde{\Sigma}))/G}{\hat{s}(\hat{\Sigma})/(NT-K)},$$

where the variance-covariance matrix for the first round estimates of the unrestricted structural model is denoted $\tilde{\Sigma}$ and second round of unrestricted sum of squares is given by $\hat{s}(\hat{\Sigma})$, the variance covariance matrix for the first round estimates of the restricted structural model is denoted $\tilde{\Sigma}$ and second round restricted sum of squares is given by $\tilde{s}(\tilde{\Sigma})$ and where the unrestricted sum of squares, based on the restricted variance covariance matrix from the first round, is denoted $\hat{s}(\tilde{\Sigma})$. N denotes the number of equations (in this case $N = 21$). K is the number of parameters for the unrestricted structural model and G is the number of restrictions.

We consider two tests concerning the nature of food demand. First we want to test the hypothesis that food expenditure is separable from non-food expenditures. That is, can we consider expenditure on food separately from expenditures on other goods? In the context of our model, this is equivalent to restricting $g_i = 0, i = 1\ldots 21$.

$\hat{s}(\hat{\Sigma})$	1211.87	Approximate F-Stat	2.3802
$\tilde{s}(\tilde{\Sigma})$	1206.49	P-Value	0.00045
$\hat{s}(\tilde{\Sigma})$	1152.78		

Thus we reject the hypothesis that expenditures on food can be considered independently of non-food expenditures. From an applied demand analysis perspective, failure to consider non-food expenditures will result in biased estimates. From a public policy perspective failing to consider non-food expenditure may overstate the effectiveness of food expenditure support programs, such as food stamps.

The second test is whether prices and demographics are sufficient to explain shifts in food demand or whether habit formation plays a role. This is important from a public policy perspective because the presence of habit formation will mitigate the impact of price changes. That is, a public policy that acts upon prices will have a lower short and intermediate run impact in the presence of habit formation. In order to test for this in our model we restrict the coefficients on lagged quantities to be equal to zero.

$\hat{s}(\hat{\Sigma})$	1211.87	Approximate F-Stat	5.9343

$\tilde{s}(\tilde{\Sigma})$	1219.58	P-Value	0.0000
$\hat{s}(\tilde{\Sigma})$	1084.99		

Thus we strongly reject the hypothesis that habit formation is absent.

8.0 Conclusion

This paper represents a first attempt at developing a method to measure the nutritional effects of policies that affect the prices of foodstuffs. In this paper we constructed a rank two demand system that nests income and prices both linearly and in logarithmic form. The model incorporated two data dependent moments of the income distribution as well as several demographic shifters. We then used this model to construct price and income elasticities for several nutrients. The next steps in this line of research are fairly clear. First, we need to estimate a full rank three demand model. Using the approach described above, our model would nest both the translog and the extended quadratic expenditure system. Secondly we need to use the full rank three demand system to test for symmetry rather than imposing it. In addition, we should test for negative semidefiniteness of the Slutsky substitution matrix. This is crucial because in ensures that price and or income elasticies of demand for food have the theoretically correct signs. Finally, we need to calculate income and price elasticities over different ranges of the income distribution, and test whether they are equal.

References

Christensen, L.R., D.W. Jorgensen and L.J. Lau. "Transcendental Logarithmic Utility Functions." *American Economic Review* 65(1975):367-383.

Gorman, W. M. "On a Class of Preference Fields." *Metroeconomica* 13 (1961): 53-56.

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

LaFrance, J. and T.K.M. Beatty. "The Structure of U.S. Food Demand." Under Revision for the *Journal of Econometrics*.

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

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

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

McDonald, J. B. "Some Generalized Functions for the Size Distribution of Income." *Econometrica* 52 (1984): 647-663.

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

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

Pollak, R.A. and T.J. Wales. “Demographic Variables in Demand Analysis.” *Econometrica* 49(1981):1533-1551.

TABLE 1.

		Mean	Standard Deviation	R-Squared	Standard Error of Regression
Fresh Milk and Cream	E1	34.50799	7.9128	0.998200	1.48692
Butter	E2	8.08872	5.43662	0.996011	1.43463
Cheese	E3	11.52309	7.21555	0.997300	1.14862
Frozen Dairy Products	E4	4.13331	1.16352	0.988523	0.41550
Canned & Powdered Milk	E5	3.19902	0.92296	0.965769	0.66005
Beef and Veal	E6	65.67647	23.43018	0.979292	10.75440
Pork	E7	34.20717	7.04967	0.941533	6.60044
Other Red Meat	E8	9.71085	2.08948	0.955603	0.95560
Fish	E9	7.81043	3.33484	0.988067	1.02970
Poultry	E10	15.67308	4.30919	0.962432	2.33133
Fresh Citrus Fruit	E11	4.41517	1.0715	0.859285	1.62735
Fresh Noncitrus Fruit	E12	11.37277	4.57526	0.976691	3.09473
Fresh Vegetables	E13	15.62052	3.84241	0.985037	1.26159
Potatoes	E14	7.95486	1.83451	0.956194	1.63702
Processed Fruit	E15	23.13193	11.59372	0.985401	4.40483
Processed Vegetables	E16	10.90067	2.70241	0.984879	1.00953
Fats and Oils	E17	13.3627	1.99345	0.964587	1.42140
Eggs	E18	15.67482	7.55406	0.997363	1.55892
Flour and Cereals	E19	19.31488	3.46391	0.996850	0.93971
Sugar	E20	25.36445	5.84785	0.988888	2.11308
Coffee, Tea & Cocoa	E21	12.10977	3.25347	0.968252	2.15249

FIGURE 1

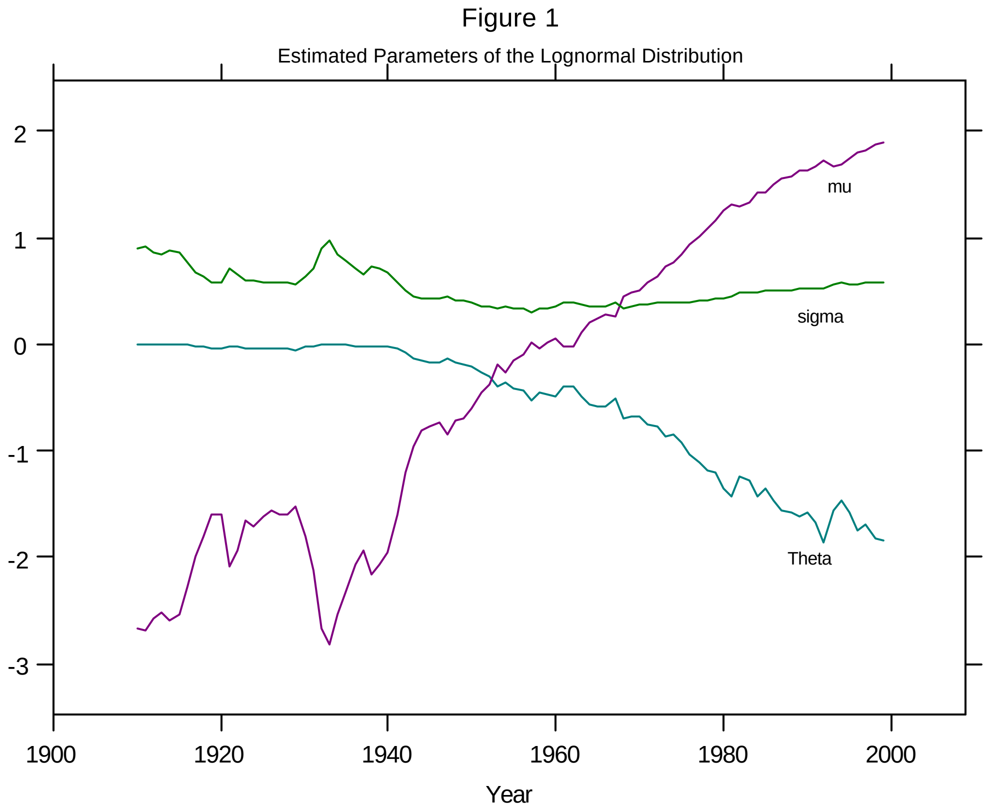


Figure 1
Estimated Parameters of the Lognormal Distribution
2
1
0
-1
-2
-3
mu
sigma
Theta
1900
1920
1940
1960
1980
2000
Year

FIGURE 2

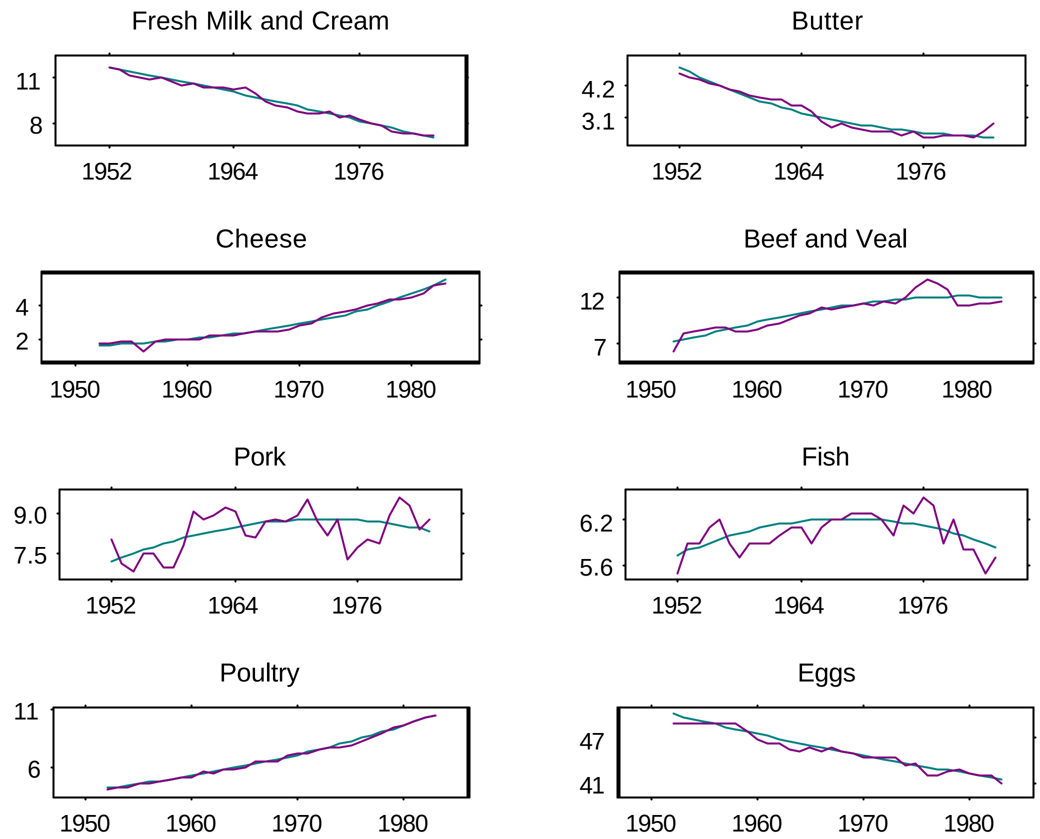
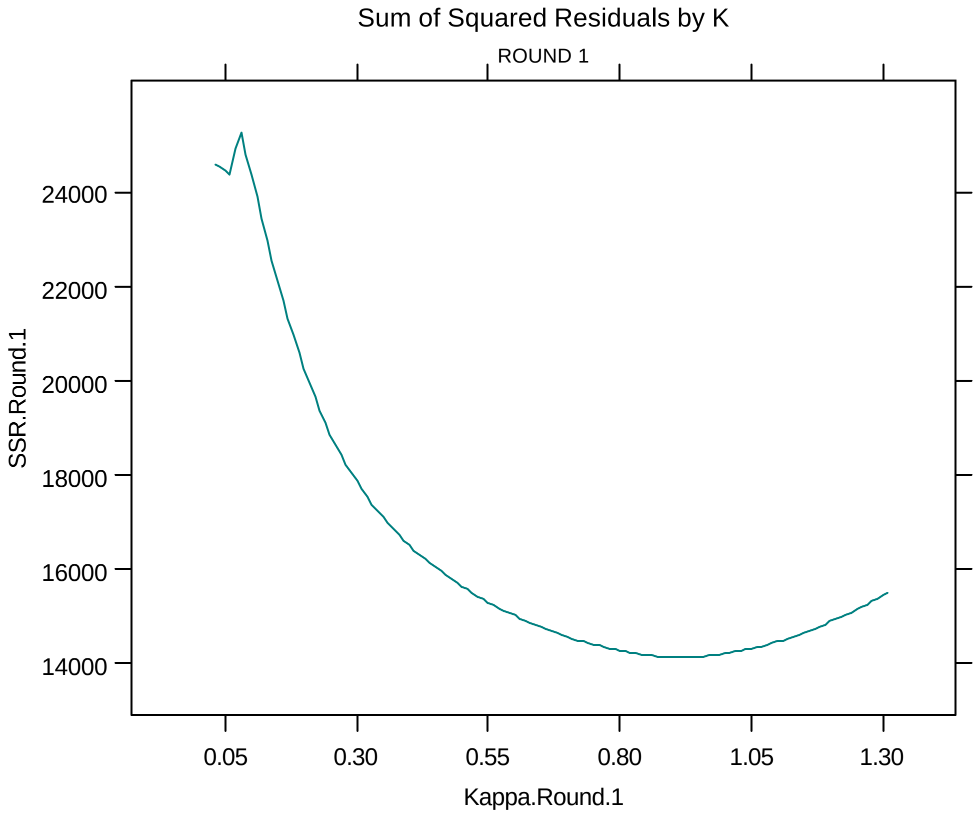

Fresh Milk and Cream
11
8
1952
1964
1976
Butter
4.2
3.1
1952
1964
1976
Cheese
4
2
1950
1960
1970
1980
Beef and Veal
12
7
1950
1960
1970
1980
Pork
9.0
7.5
1952
1964
1976
Fish
6.2
5.6
1952
1964
1976
Poultry
11
6
1950
1960
1970
1980
Eggs
47
41
1950
1960
1970
1980

FIGURE 3.


Sum of Squared Residuals by K
ROUND 1
24000
22000
20000
18000
16000
14000
SSR.Round.1
0.05
0.30
0.55
0.80
1.05
1.30
Kappa.Round.1

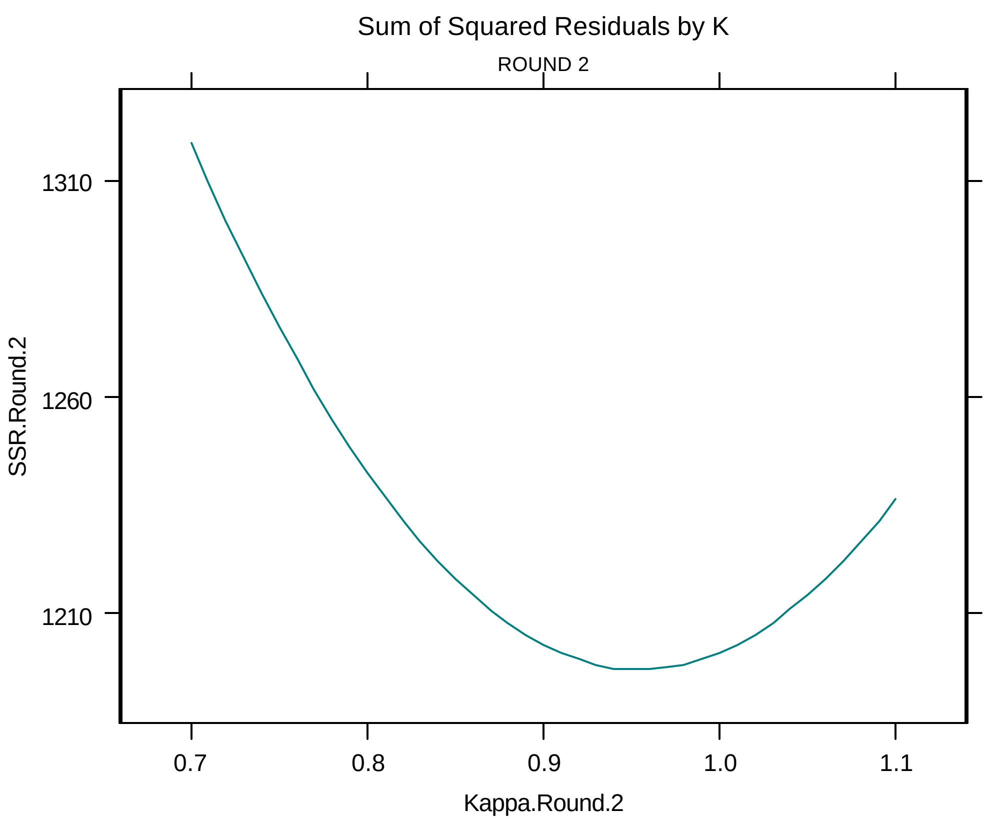
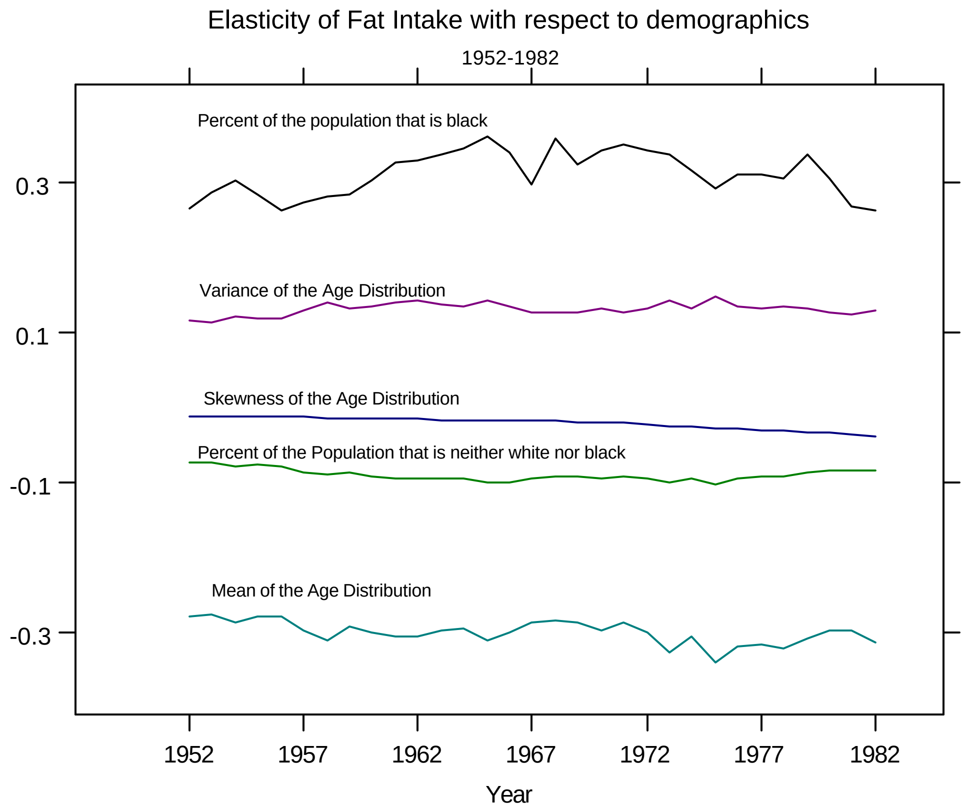

Sum of Squared Residuals by K
ROUND 2
1310
1260
1210
SSR.Round.2
0.7
0.8
0.9
1.0
1.1
Kappa.Round.2

Figure 4.



Elasticity of Fat Intake with respect to demographics
1952-1982
Percent of the population that is black
Variance of the Age Distribution
Skewness of the Age Distribution
Percent of the Population that is neither white nor black
Mean of the Age Distribution
0.3
0.1
-0.1
-0.3
1952
1957
1962
1967
1972
1977
1982
Year

Figure 5.

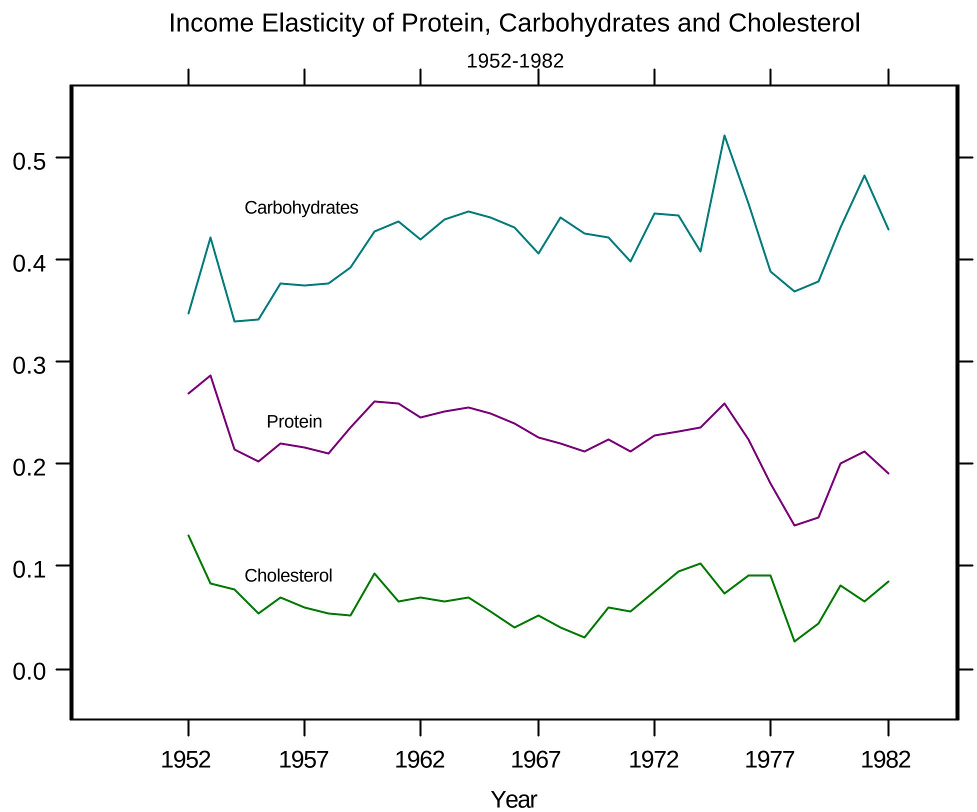
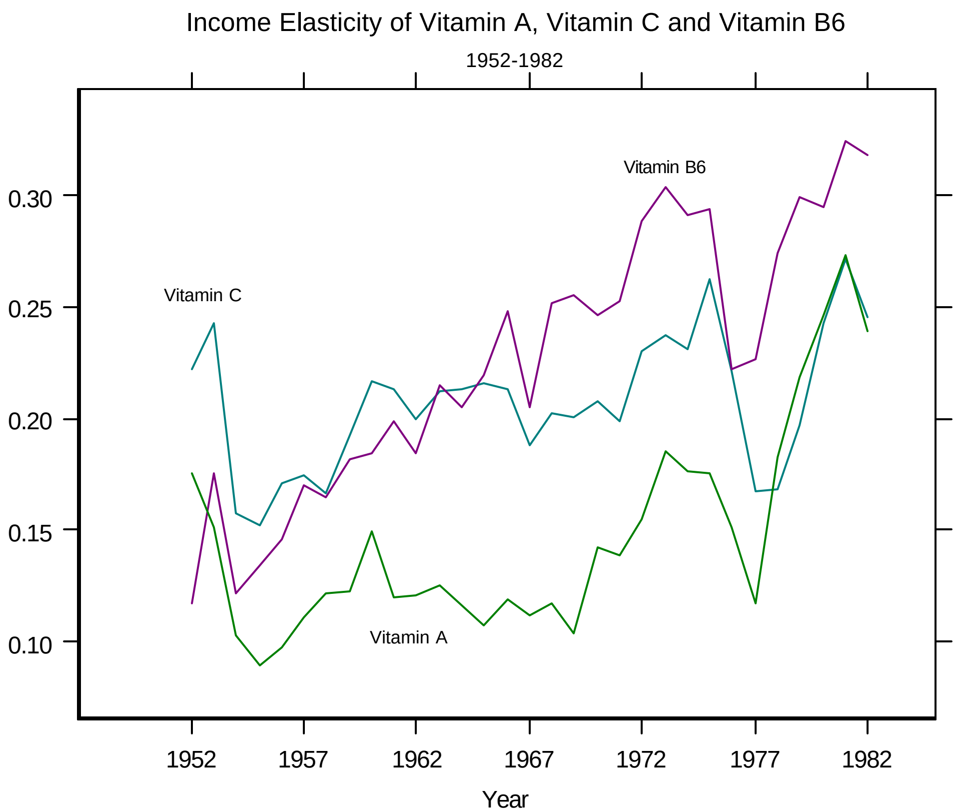

Income Elasticity of Protein, Carbohydrates and Cholesterol
1952-1982
Carbohydrates
Protein
Cholesterol
0.5
0.4
0.3
0.2
0.1
0.0
1952
1957
1962
1967
1972
1977
1982
Year

Figure 6



Income Elasticity of Vitamin A, Vitamin C and Vitamin B6
1952-1982
Vitamin B6
Vitamin C
Vitamin A
0.30
0.25
0.20
0.15
0.10
1952
1957
1962
1967
1972
1977
1982
Year

Figure 7.

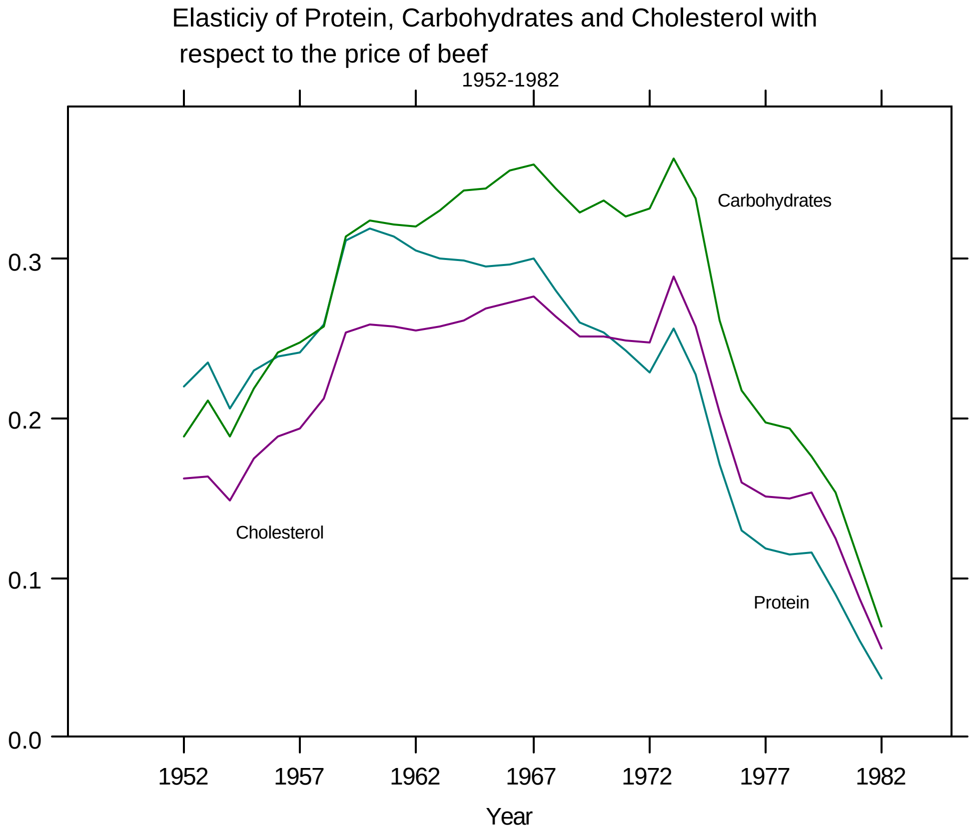

Elasticiy of Protein, Carbohydrates and Cholesterol with respect to the price of beef
1952-1982
Carbohydrates
Cholesterol
Protein
0.0
0.1
0.2
0.3
1952
1957
1962
1967
1972
1977
1982
Year

Figure 8.

Elasticity of Vitamin A, Vitamin C and Vitamin B6 with respect to the price of fresh vegetables

1952-1982