Working Papers in
AGRICULTURAL ECONOMICS
October 1992 92-11

INCORPORATING DEMOGRAPHIC VARIABLES INTO DEMAND SYSTEMS FOR FOOD: THEORY AND PRACTICE

Deborah C. Peterson
and
Timothy D. Mount

Department of Agricultural Economics
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853-7801

It is the policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

INCORPORATING DEMOGRAPHIC VARIABLES INTO DEMAND SYSTEMS FOR FOOD: THEORY AND PRACTICE

Deborah C. Peterson
Economist
Australian Bureau of Agricultural and Resource Economics
Canberra, Australia

Timothy D. Mount
Professor
Department of Agricultural Economics
Cornell University
Ithaca, NY

Abstract

The primary focus of the theoretical literature on incorporating demographic variables into demand models is on modifying the form of the utility function. This approach, however, tends to impose restrictive implications about behavior. Consequently, modifying the demand equations using household equivalent scales is the procedure adopted. Continuous scales for age are used together with variables representing race, life style and employment status. The model is fitted to data from the Consumer Expenditure Diary Surveys for 1980 and 1986 for nine food groups in the US. Demographic variables are shown to be particularly important for explaining expenditure on food away from home. Tests for homogeneity across years, seasons and regions are rejected implying that the effects of demographic variables vary within the sample.

Introduction

Many studies support the notion that demographic variables are important determinants of demand for food products in the United States: for recent examples, see Blaylock and Smallwood (1986), Buse et al (1987), Gould, Cox and Perali (1991); Heien and Pompelli (1988), McCracken and Brandt (1987); Price (1986), and Schrimper (1986). Consequently, ignoring the effects of demographic variables on the demand for food constitutes a serious misspecification of a demand system. Nevertheless, there have been relatively few studies of demand for disaggregated food products in the United States within the framework of complete demand systems which have employed micro-level data and which also take into account the effects of demographic variables. A notable exception is a study by Kokoski (1986), which examines the demand for five groups of food-at-home using a Quadratic Expenditure System.

An important issue for modeling is that the theoretical basis for incorporating demographic variables is limited. The purpose of this research is 1) to review the theoretical basis for incorporating demographic variables into demand systems, 2) to specify an improved model, and 3) to estimate demand for nine exhaustive disaggregated food groups: cereals and bakery products (cerbak); meats, poultry, fish and eggs (mpfe); dairy products (dairy); fruits and vegetables (frutveg); sweets; fats and oils (fats); nonalcoholic beverages (nonalbev); food-away-from-home (fafh); and miscellaneous prepared foods. The Almost Ideal Demand System developed by Deaton and Muellbauer (DM) is used because it has a number of appealing characteristics for modeling food demand. In particular, it imposes no a priori restrictions on inter-commodity substitution characteristics (Deaton and Muellbauer 1980a). The specification of demographic variables incorporates Tyrrell's continuous household composition scale for age and sex to form part of a household equivalence scale (Tyrrell 1979, Tyrrell and Mount 1987). In addition, household characteristics are included as discrete variables to account for factors such as employment status. These demographic variables are particularly important for explaining the demand for food-away-from-home. Data are from the 1986 Consumer Expenditure Diary Survey (CEDS).

Alternative theoretical models for demographic variables

While it is clear that demographic variables have an important impact on household consumption patterns, it is not at all clear how these effects should be modeled. All standard procedures, including the Engel (Engel 1895), Prais-Houthakker (Prais and Houthakker 1955), modified Prais-Houthakker (Pollak and Wales 1981), Barten Scaling (Barten 1964), Translating (Pollak and Wales 1978), Gorman and Reverse Gorman (Gorman 1976, 1981) procedures, scale quantities in the utility or demand functions by some function of demographic characteristics, i.e., an equivalence scale.

Equivalence scales essentially define household size as a weighted function of the number of individuals of different types in the household (Deaton and Muellbauer 1980a). Although originally developed as an improvement on per capita models, Deaton and Muellbauer (1980a) note that an equivalence scale can more generally be considered a "constant utility, constant price, cost-of-living index relating the cost of living at household characteristics A to that of some reference household with characteristics A^R." The use of an equivalence scale is not restricted to deflation of total expenditure, but rather can be seen as a parameter of the cost function. In a sense, any function of household demographic variables could be considered an equivalence scale if this function is normalized for some reference household.

All of the standard procedures for defining equivalence scales involve some restriction of the way demographic variables enter the demand system. For example, if the Prais-Houthakker model is considered in the context of utility maximization, Muellbauer (1974, 1980) notes that the absence of relative price effects in the model implies that all goods are consumed in fixed proportions. In other words, no substitution between goods is permitted, and inferior goods and the possibility that the demand for any good reaches a plateau are ruled out. Engel's model, a special case of the Prais-Houthakker model where the function of household characteristics is the same for all goods, is even more restrictive, since, for example, the needs of children relative to adults are assumed to be the same for all goods.

Pollak and Wales modified the Prais-Houthakker model to allow substitution effects, but the resulting demand system is consistent with utility maximization if and only if the underlying utility function is additively separable (Pollak and Wales 1981, Lewbel 1986).

Barten Scaling, Translating, and the Gorman and Reverse Gorman procedures are all based in the utility maximization framework and involve scaling the quantity arguments in the utility function by some function of demographic characteristics. For the procedures beginning with the utility function, the exact effect of the transformation depends on the choice of functional form for the utility function and the resulting form of the demand system (Pollak and Wales 1981). For example, when translating is applied to any demand system that is linear in expenditure, marginal budget shares do not vary with the demographic variables (Pollak and Wales 1981).

Unfortunately, as noted by many authors (see, for example, Muellbauer 1977, Deaton and Muellbauer 1980a, and Nelson 1987), the quasi-price effect of demographic variables, which is the essence of the Barten procedure, can lead to some paradoxical results. To give the classic example, suppose only children drink milk, and only adults drink beer. Under Barten scaling, the net response to the addition of a child to the family could be a decrease in milk consumption and an increase in beer consumption. This example points out that the substitution effects predicted by the model may be excessive, and this has been a major source of criticism of the Barten procedure. The translation term in the Gorman and Reverse Gorman procedures tend to reduce the possibilities of excessive substitution found in Barten scaling (Muellbauer 1977). However, both procedures still result in demographic variables entering the utility function in a restrictive way.

The utility-based procedures discussed above assume that demographic variables alter preferences in a specific manner. An advantage is that these procedures are universally applicable, in the sense that they may be applied to any form of utility function. The disadvantage is that the possible range of demographic effects that can be achieved by these procedures is limited (Lewbel 1985).

An alternative to procedures involving a transformation of the utility function is to directly allow some of the parameters of a demand system to vary with demographic variables. This is the implicit result of ad hoc models which add demographic variables to the right hand side of a regression equation. The advantage of such an approach is the great flexibility possible in terms of the range of demographic effects that can be achieved. The disadvantages are its ad hoc nature, and, as Lewbel (1985) notes, its lack of "general applicability, being specific to the given starting model."

An unpooled specification, which involves separate analysis of groups of households which are relatively homogeneous in the salient demographic characteristics, might be preferred, since it does not involve parameter restrictions. However, other important aspects to be considered are data availability and the number of parameters to be estimated. As the number of demographic variables and the number of categories within each variable increase, the unpooled specification quickly becomes very expensive in terms of degrees of freedom. Also, if there are functional relationships between the unpooled groups, this information will be lost if an unpooled specification is used.

Lewbel (1985) presents a general method of introducing demographic effects into a demand system through the use of "modifying functions" that modify the system's cost function while preserving the properties of the unmodified cost function that make it admissible as a cost function. Lewbel's modifying function approach has the universal applicability of traditional utility transforming procedures, since the modifying functions can be applied to any cost function. It also has the advantage of the individual model modification approach of allowing interaction of demographic variables with prices and expenditure in an almost unlimited number of ways. The modifying function method is shown to encompass the Engel model, Barten Scaling, Translating, Gorman and Reverse Gorman procedures. The approach of directly allowing some parameters in a demand system to vary with demographics is also shown to be possible through modifying functions, although the form of the modifying functions might be complicated.

Through the modifying function approach, Lewbel extends the Barten Scaling and Translating procedures to budget share equations. Demographic budget share translating is defined as a procedure for allowing a constant term in a budget share equation to be a function of demographic variables. For example, allowing the constant term in the DM model to vary with demographic variables would constitute demographic translating according to Lewbel's scheme. Demographic budget share scaling is defined analogously as deflating budget shares by adult equivalent scales. Presumably a Gorman or Reverse Gorman budget share procedure could also be defined as combinations of budget share scaling and translating.

However, none of the procedures mentioned above identify which demographic variables affect utility or demand or how these effects should be specified. This is because there is no attempt to model behavior of a group of individuals. Furthermore, conclusions of studies comparing different approaches (Muellbauer 1977, Pollak and Wales 1981, and Barnes and Gillingham 1984) are based on statistical fit rather than tests of theory. Until such theory has been developed, incorporating demographic variables into a demand system remains largely an empirical matter. Current theory places almost no restrictions on how this should be done.

The specification of equivalence scales

There are a number of ways such scales may be specified. The alternatives may be classified as either discrete or continuous. In the case of discrete scales, the value of the scale is assumed to be constant within defined categories of the demographic characteristics. Continuous scales, on the other hand, take advantage of the continuous nature of the demographic characteristic age.

There are two major limitations of the discrete approach. The first is the arbitrary nature of the choice of categories. This is a problem because the estimated values of the scale can be sensitive to the boundaries of the categories. The second is the discontinuity of the scale values from one group to the next and the lack of a systematic relationship between the scale values of each group. The behavior of households or individuals on either side of a boundary is more likely to be similar to each other than to households or individuals on the opposite side of each category. For continuous equivalence scales, all scale values are systematically linked by a defined functional relationship. The parameters of a continuous scale are the coefficients of the function rather than the values of the scales.

Two popular specifications of a continuous equivalence scale are those of Blokland (1976), and Buse and Salathe (1978). Blokland chooses two cubic functions in age, one for males and one for females under the age of 20 years, and two constant functions for males and females over the age of 20. Restrictions are imposed to force continuity at the boundary of age 20. Buse and Salathe modify the Blokland approach: instead of using constant functions for males and females after the age of 20, their scale only remains constant from age 20 to 55 and for ages over 75. The scale follows another cubic function from ages 55 to 75. Tedford, Capps, and Havlicek (1984) follow the Buse and Salathe approach in their study of convenience foods in the United States.

Selection of boundary values of the age classes becomes an issue when different functions are spliced, and parameter estimates may be sensitive to these boundary values. Tedford, Capps, and Havlicek (1986) note that previous work provides no justification for the selection of boundary values or what causes consumer behavior to be different at various stages. The authors address this problem by using the concept of the family life-cycle from the fields of psychology and human development to develop the scales. Four developmental periods are identified: infancy, childhood and adolescence (birth to 17 years of age), early adulthood (22 to 40 years of age), middle adulthood (45 to 60 years of age), and late adulthood (65 to 80 years of age). Each of the remaining five time periods are referred to as transitional stages. Tedford et al allow four different cubic functions for each sex for each of these phases.

Constant functions are assumed for each sex for each transitional age and for all ages over 80.

They compare this procedure with the Blokland et al models for expenditures on food at home. The results suggest that Blokland's model is too restrictive to adequately portray expenditures over the household life-cycle. However, the empirical results are found to be similar to those obtained by the Buse and Salathe model. Blokland at al, and Tedford et al all use single equation linear Engel functions.

Tyrrell avoids the difficulty of choosing boundary values by specifying a commodity-specific household composition scale as the exponential of the sum of individual equivalence scales that are single cubic polynomials in age for each sex. Tyrrell then specifies a household equivalence scale as the product of actual household size and the household composition scale. The interaction between household size and composition allows the marginal effects of household members to vary according to household size. This is not permitted in the models of Blokland et al, and Tedford et al, making these models more restrictive in that respect.

The model presented in the next section follows Tyrrell's approach and extends it by adding other demographic variables such as the number of full- and part-time workers, household type and retirement status. The change in food consumption behavior around age 65, apparent in the work by Tedford et al (1986) and Buse and Salathe, is probably due to retirement from work and associated changes in lifestyle rather than the effects of reaching age 65 per se. Retirement is a major milestone in an income-earner's life. It is often associated with an abrupt change of lifestyle, such as moving to a new (often warmer) region. These variables are shown to be particularly important for explaining expenditures on food away from home. The range of demographic characteristics considered in many studies has been relatively narrow. Many applications are limited by the available data, but the data set used in this study is rich in demographic information. Three types of demographic variables are incorporated into the equivalence scale: 1) member characteristics such as the age and sex of household members; 2) household characteristics such as household size and type; and 3) locational characteristics such as season and region.

Following Tyrrell, a household equivalence scale, S, includes an exponential sum of cubic polynomial functions of the log of age and sex of household members. An exponential form gives positivity, and has the advantage that log(S) is linear in parameters for ease of estimation. The log of age is used rather than age for two reasons (Tyrrell 1979): first, the logarithmic transformation accentuates the differences between, say, five and ten year-olds as compared with the difference between 50 and 55 year-olds, which is reasonable since the most dramatic changes in food consumption take place in the early years of life; and second, this transformation dampens potential explosiveness of the cubic polynomials at high ages.[1]

For convenience, the cubic functions are expressed in terms of four Lagrange interpolation polynomials L_1, L_2, L_3, and L_4, corresponding to four chosen reference ages. Two of the four reference ages, 0 and 100, are chosen to include endpoints corresponding to birth and an extremely long life. The other two reference ages, 14 and 20, are chosen to coincide with probable

[1] So that a very young child (one whose age is close to zero) does not cause too large of an effect, 1.5 is added to all ages. In the CEDS, the age of children under one year old is recorded as zero, and then adding some small constant becomes a necessity, as the logarithm of zero is undefined. Choosing the constant to be different than one ensures that a household with, say, all males and no females is different than a household with all males and a female child under one year in age.

inflection points of the polynomial corresponding to the beginning of adolescence and adulthood. The choice of reference ages affects details of the calculations and interpretation of the regression coefficients, but not the final results.

Age and sex of household members are the only member characteristics included in the model. Other factors such as race and education are treated as household characteristics to save degrees of freedom. A single index variable is created to account for the effect on household consumption of each member characteristic treated in this manner. For example, race of the reference person is used as an indicator of the racial composition of the household. The error associated with this is likely to be small, since less than two percent of households are of mixed race.

The complete specification of the household equivalence scale for commodity n is

$$S_{nh} = Z^{d_n} \exp[\Sigma_{males} f_{nm}(age) + \Sigma_{females} f_{nf}(age)] \exp(g_{n1}D_1 + ... + g_{nJ}D_J)$$

where Z is household size for household h; f_{nm} and f_{nf} are the cubic weighting functions of age for males and females, respectively; and D_i (i = 1,2,..., J) are dummy variables for other household characteristics.

Specification of the demand system

The usual form of the DM model expresses budget shares as a linear function of the log of prices and real income. Prices are not available on the public use tapes of the data used in this study, and consequently regional and seasonal dummy variables are used to represent groups facing homogeneous prices. These dummy variables will also account for lifestyle effects associated with living in different regions and seasons. This method of accounting for prices is less than ideal. It would be preferable to have more detailed locational information available to identify groups of households with regional price indices, such as those for Standard Metropolitan Statistical Area (SMSA). Unfortunately, the CEDS public use data does not contain such information due to non-disclosure requirements.

Assuming prices are constant within regions and seasons in a cross-section analysis, and scaling income by some commodity-specific function of demographic variables, the DM system collapses to a function of nominal income adjusted by the household equivalence scale. The resulting model is:

$$
\begin{aligned}
w_{nh} &= a_n + b_n \log M_h + c_n \log S_{nh} \\
&= a_n + b_n \log M_h \\
&\quad + b_{n1}\Sigma_{males} L_1 + b_{n2}\Sigma_{males} L_2 + b_{n3}\Sigma_{males} L_3 + b_{n4}\Sigma_{males} L_4 \\
&\quad + b_{n5}\Sigma_{females} L_1 + b_{n6}\Sigma_{females} L_2 + b_{n7}\Sigma_{females} L_3 \\
&\quad + b_{n8}\Sigma_{females} L_4 + d_n \log Z + g_{n1} D_1 + \ldots + g_{nJ} D_J
\end{aligned}
$$

where w_{nh} is the budget share of good n for household h, M_h is income, S_{nh} is the household equivalence scale, $L_1, \ldots, L_4$ and the Lagrangian interpolation polynomials, b_n and c_n are commodity-specific parameters, and a_n is a parameter which absorbs the price variables. Maintaining adding-up ($\Sigma_n W_{nh} = 1$) requires that:

$$\Sigma_n b_n = \Sigma_n d_n = 0;$$

$$\Sigma_n g_{nj} = 0 \qquad (j = 1, 2, \ldots, J);$$

and

$$\Sigma_n b_{nk} = 0 \qquad (k = 1, 2, \ldots, 8).$$

The expenditure elasticity for good n is given by $1 + b_n/w_{nh}$. It can be seen that demographic variables can influence both the size and sign of the expenditure elasticity, and consequently, help to determine the classification of goods as luxuries, necessities, or inferior goods.

A two-stage budgeting model is used, where households first allocate total expenditures between food and nonfood, and then allocate total expenditure on food among food groups. The expenditure variable in the first stage of the budgeting process is problematic. Ideally, the sum of food and nonfood expenditures should be used, but this unfortunately is not available. Instead, a proxy variable, after-tax income, is used for this purpose.

Data

Data for the CEDS are collected from a nationwide probability sample of households designed to represent the total civilian non institutional population. Consumer units are asked to keep a detailed daily diary of their expenses for two consecutive one-week periods. Only expenditures, not prices and quantities purchased, are available on the public use tapes for these surveys. For this study, reported expenditures for each week that a consumer unit participated in the survey are averaged, so that the level of analysis is weekly household consumption. The expenditure data are adjusted to December 1986 dollars using regional monthly price indices based on the Consumer Price Index for All Urban Consumers.

The sample selected is urban households inside Standard Metropolitan Statistical Areas (SMSAs) whose weekly reported total food expenditure is greater than zero. Food items bought as gifts for people outside the consumer unit are removed from the reported expenditures, since the goal of this research is to explain expenditure on items purchased by the consumer unit for its own consumption. These restrictions resulted in a sample size of 5252 households.

Allocation of total food expenditures among disaggregated food groups is the second stage of the budgeting procedure. No sample restrictions other than those already described are necessary for estimation of this stage.

The following additional sample restrictions are made for stage one estimation. Households with unreliable income data, as indicated by the BLS, and those with household type "other" were eliminated.[2] Since expenditures can never be negative, households whose income after-tax is less than zero are also eliminated. Finally, since after-tax income is being used as a proxy for total expenditures, households who reported spending more on weekly total food expenditures than their weekly after-tax income are eliminated. After these eliminations, 3719 urban inside-SMSA households remain for inclusion in the stage one sample.

Weekly per capita income and food expenditure by the stage one sample is $25.59, while weekly per capita income after tax is $199.03 (with standard deviations of $28.39 and $262.58 respectively).[3] Budget share sample means are presented in Table 1. The largest proportion of the total food budget is spent on food away from home (37 percent). Expenditure on meats, poultry, fish and eggs account for 17.5 per cent of total food expenditure. Fruits and vegetables, cereals and dairy products each account for about 10 percent of total food expenditure. Sweets, fats and oils, and nonalcoholic beverages each

---

[2] Households with household type "other" were eliminated since this group is composed of non-homogeneous households whose food purchasing patterns are likely to have a very different structures.
[3] 1986 dollars, N = 9434 (stage one sample).

account for only a small proportion of food expenditure. Other sample statistics are presented in *author publication* (1991).

Table 1 Budget share sample means

Expenditure category	Mean	Std. dev.
Stage one	(N = 3719)	
Food	0.175	0.144
Nonfood	0.825	0.144
Stage two	(N = 5252)	
Cereals and bakery products	0.092	0.075
Meats, poultry, fish, and eggs	0.175	0.140
Dairy products	0.086	0.077
Fruits and vegetables	0.107	0.090
Sweets	0.020	0.030
Fats and oils	0.016	0.024
Nonalcoholic beverages	0.062	0.071
Miscellaneous foods	0.072	0.077
Food away from home	0.370	0.286

The definitions of demographic variables

The selected member characteristics are age and sex.[4] The selection of household characteristics is based on the discussion in Chapter 10 of Food Demand Analysis (Raunikar and Huang 1987). Six household characteristics used in the analysis are household size; race; retirement; education; household type; and the number of full- and part-time income earners.

[4] In the 1986 survey, if the age of an individual is greater than 90, the value of the variable 'age' is set to 90 and flagged as topcoded. There are 29 out of 14141 members of urban inside-SMSA households (0.2 %) whose age is topcoded. For these cases, 90 was taken to be the age of the household member. A similar approach was followed with topcoded income.

Household size is defined as an integer variable. Race is a dummy variable for a white or nonwhite reference person. Retirement is treated in a similar way. Education is a categorical variable which equals one if neither the reference person nor the spouse has any college education and zero otherwise.

Household type is used to capture life-style effects using the following eight BLS-defined categories:

1. Husband and wife only.
2. Husband and wife, own children only, oldest child less than six.
3. Husband and wife, own children only, oldest child older than 6 but younger than 18.
4. Husband and wife, own children only, oldest child older than 17.
5. All other husband and wife consumer units.
6. Single parent, with own children, at least one child is less than 18.
7. Single consumers.
8. Other consumer units.

It is hypothesized that life-style is also affected by the number of full- and part-time income earners in the household. For example, it might be expected that a couple with full-time jobs might prefer less time-intensive meal preparation than another couple with the same income level, but with only one earner. Six dummy variables are created corresponding to the number of full- and part-time workers in the household. Full-time workers are defined as those who usually work 35 hours or more per week. The categories are:

1. Zero earners.
2. One full-time, zero part-time earners.
3. Two or more full-time, zero part-time earners.
4. One full-time, one part-time earner.

5. Zero full-time, one part-time earner.

6. Other full-time/part-time combinations.

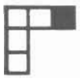
For the locational characteristics there are four seasons and four regions (Northeast, Midwest, South, and West). Dummy variables are included to represent the assumed homogeneity in prices within each season-region. However, locational characteristics are also hypothesized to account for difference in taste between season-regions, so separate models are estimated for each season-region to allow for interactions with other model parameters. As shown in the next section, taste differences cannot be accounted for by simple intercept differences.

Estimation and results

A stochastic specification for the demand system is obtained by adding a random error term to each equation. For complete systems of demand equations such as the DM model, the error variance-covariance matrix is singular because the budget shares sum to one by construction. The usual procedure of dropping one good (nonfood in stage one and miscellaneous foods in stage two) from the system and estimating by Zellner's Iterative Seemingly Unrelated Regression (ITSUR) is followed, which is asymptotically equivalent to a maximum likelihood procedure. Since there are no cross-equation constraints, the estimates are identical to ordinary least squares. However, the estimated variance matrix of residuals across equations is required for testing purposes.

The data are randomly divided into two sub-samples. One half is used for exploratory analysis, while the other half is retained for testing. Testing of the model is performed with two objectives. First, to examine whether the season/region structure of the model can be simplified. Second, to test whether or not demographic variables are important. The sequence of model estimation and testing for both the stage one and stage two models is to first consider simplification of the season-region structure, and then to consider reducing the number of demographic variables in the model. This strategy is summarized in Table 2, where an entry indicates that a model is estimated with the indicated features. This table also introduces icons which are later used to identify the models.

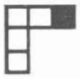
Table 2 Modeling strategy

	Season and region affect all model parameters	Season affects the intercept, region affects all model parameters	Season and region affect only the intercept
All demographic variables included in model			
Only household size included in model			
No demographic variables included in model			

Model coefficients are not presented here due to space constraints. Parameter estimates for the unrestricted model for all season-regions for each good are presented in *author publication* (1991); parameter estimates for the other models are available from the authors on request.

R^2 statistics are provided in Table 3 as a summary measure of goodness of fit. The R^2 values for the stage one model with all demographic variables included and with season and region affecting all model parameters is high for cross-sectional estimation (0.39). The R^2 value decreases considerably as the model structure is simplified. The R^2 values for the stage two models are lower than those for the stage one model, but follow the same pattern as in the first stage: they decrease quite substantially as the number of parameters is reduced.

Table 3 R-squared values

Model	Stage one	Stage two							
	Food	cerbak	mpfe	dairy	frutveg	sweets	oils	nonalbev	fafh
	0.36	0.11	0.12	0.12	0.11	0.02	0.04	0.02	0.18
	0.38	0.12	0.14	0.13	0.12	0.05	0.05	0.04	0.20
	0.39	0.17	0.20	0.19	0.18	0.10	0.10	0.09	0.24
	0.36	0.06	0.08	0.09	0.02	0.03	0.02	0.02	0.08
	0.29	0.02	0.03	0.04	0.02	0.02	0.01	0.02	0.02

Model Specification Tests

Simplification of the season/region structure and the importance of including demographic variables in the model are tested using a likelihood ratio test. The test statistic, $-2\log(L_R/L_F)$, where L_R represents the value of the likelihood function for the restricted model, and L_F represents the same for the full model, is asymptotically distributed as χ_m^2, where m is the number of restrictions needed to define the null hypothesis (Harvey 1981).

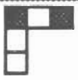
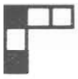
Tests of pooling across season/region groups involve the stage one and stage two models with all demographic variables included in the model. First, models are estimated with season and region affecting all model parameters. These models are then tested against models with season and region only affecting the intercept, and against models with season affecting the intercept and region affecting all parameters.

The results for both stage one and stage two, shown in Table 4 (rows 1,2 5 and 6), indicate that there are significant differences among parameters across season and regions, so the models cannot be simplified with regard to their season/region structure.

Tests of simplification of the model structure with respect to the demographic variables involve the stage one and stage two models with season and region affecting all model parameters. These models are tested against models with no demographic variables included, and against models with household size as the only demographic variable included. Household size is chosen as an intermediate step between the other two models because it is common practice to estimate models using household data by expressing total expenditures on a per capita basis.

The results for both stage one and stage two, shown in Table 4 (rows 3,4,7, and 8), indicate that all of the demographic variables should be included in the model.

Table 4 Results of the likelihood ratio tests

Models	Test statistic	Degrees of freedom	p-value	Reject the restrictions[a]
		Stage one		
	516.8	384	0.0000	√
	386.0	300	0.0000	√
	535.8	352	0.0000	√
	906.4	368	0.0000	√
		Stage two		
	5782.6	3612	0.0000	√
	4291.9	2832	0.0000	√
	4816.3	2944	0.0000	√
	5457.5	3072	0.0000	√

a. Using a significance level of 0.05.

Elasticity estimates

The income elasticity of food at the mean for the unrestricted model (with all demographic variables included and with season and region affecting all model parameters) is 0.27.[5] Income and food expenditure elasticities for disaggregated foods are presented in Table 5. Dairy products are found to be the least sensitive to changes in income. The most sensitive are the meats, poultry, fish, and eggs group, food away from home, fruits and vegetables, and fats and oils. Since the importance of the income effect on household expenditure patterns depends on the size of the budget share as well as the size of the income effect, the response for fats and oils is much less significant than the other three (see Table 1).

Table 5 Estimated elasticities for disaggregated foods[a]

Food group	Food expenditure elasticity	Income elasticity
Cereals and bakery products	0.88	0.24
Meats, poultry, fish and eggs	1.05	0.28
Dairy	0.74	0.20
Fruits and vegetables	0.97	0.26
Sweets	0.98	0.26
Fats and oils	1.05	0.28
Nonalcoholic beverages	0.92	0.25
Other	1.07	0.29
Food away from home	1.07	0.29

a. Evaluated at the mean of the data of the stage two sample.

5 Evaluated at the mean of the stage one data.

Effect of demographic variables on expenditure patterns

To illustrate the effects of demographic variables and income on household food expenditures, nine household profiles are created which represent households at different stages of a typical household life-cycle. The stage one and stage two models are used in sequence to estimate budget shares for food and disaggregated foods for these households. For the purposes of this example, model parameters are averaged over seasons and regions, as an alternative to arbitrarily selecting a particular season and region or presenting an unwieldy sixteen sets of predictions for each year.

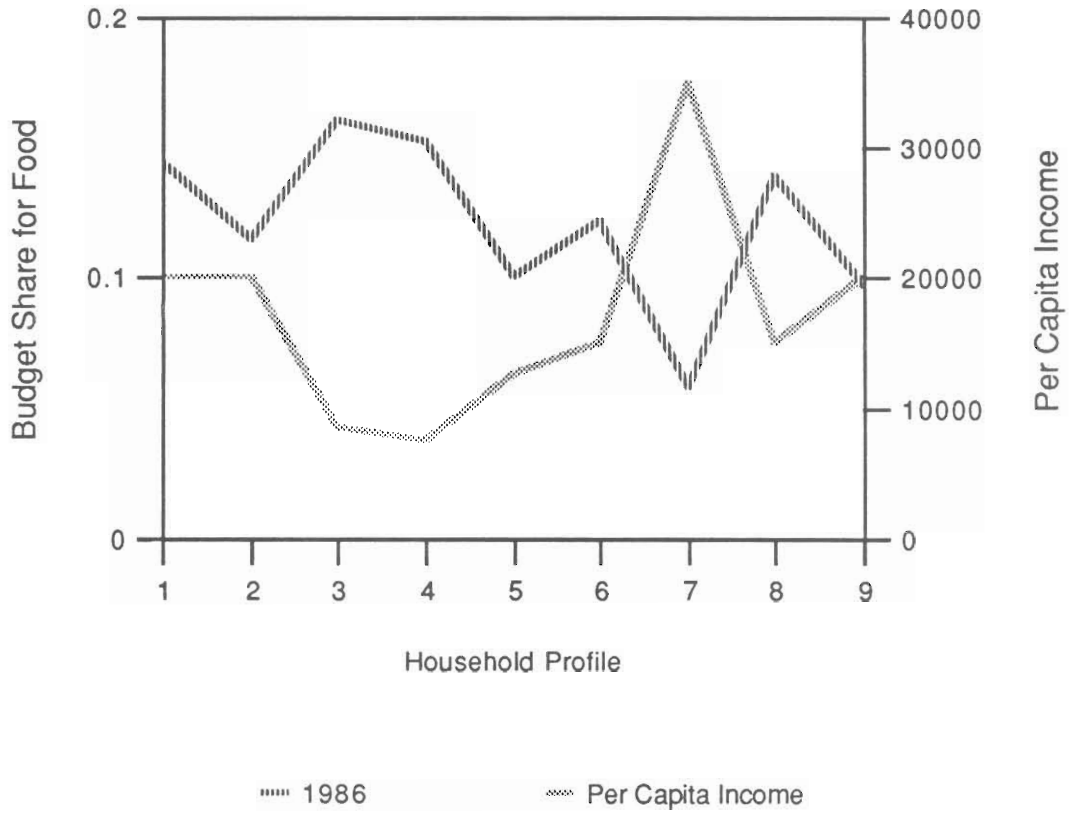
The complete specification of the profile household characteristics are given in Table 6. Household before-tax income is used as a starting point, and after-tax income is estimated using the 1986 marginal tax rates, adjusted by information obtained from the relationship between before- and after-tax income observed in the data.[6] The budget share patterns for food for each household predicted by the model are shown in Figure 1. Per capita household income is also plotted on this figure.

[6] Details of this estimation procedure are reported in *author publication* 1991.

Table 6. Profile household characteristics

Life-cycle stage	Household composition	Age of household members		Number of earners		Household before-tax income ($)
		males	females	full-time	part-time	
1	single male	22		1	0	20000
2	husband, wife	26	22	2	0	40000
3	husband, wife, one child	28, 1	24	1	0	26000
4	husband, wife, two children	31, 4	27, 1	1	0	30000
5	husband, wife, two children	36, 9	32, 6	1	1	50000
6	husband, wife, two children	44, 17	40, 14	2	0	60000
7	husband, wife	52	48	2	0	70000
8	husband, wife	66	54	0	0	30000
9	widow		80	0	0	20000

Figure 1 Food budget shares and income for profile households



Description of households:

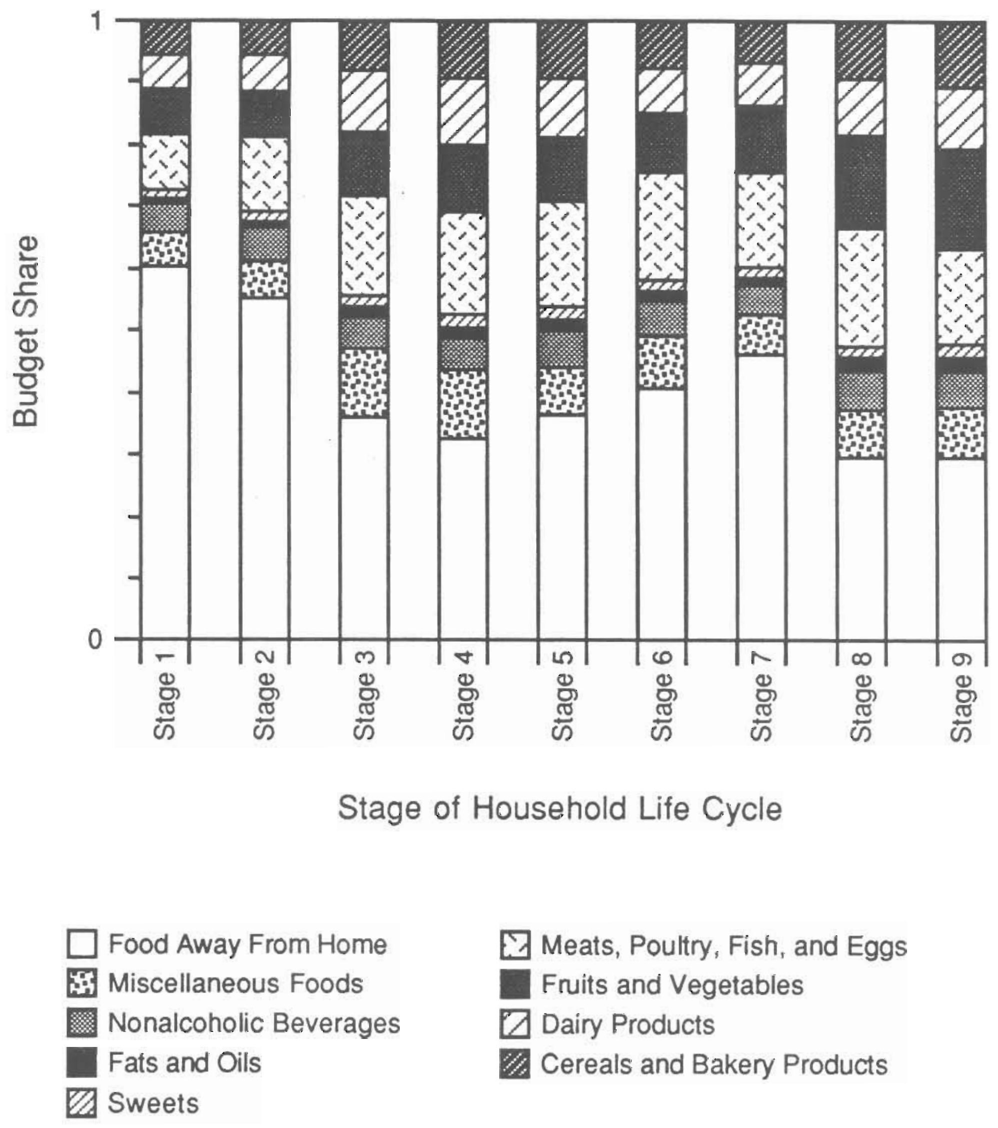
1. Single male; income before tax = $20000
2. Married couple; both employed full-time; income before tax = $40000
3. First child; wife leaves work force; income before tax = $26000
4. Second child; income before tax = $30000
5. Both children in school; wife returns to part-time work; income before tax = $50000
6. Teenage children; wife returns to full-time work; income before tax = $60000
7. Empty nest; both employed full-time; income before tax = $70000
8. Retirement; income before tax = $30000
9. Elderly widow; income before tax = $20000

In general, there is a rough inverse correlation between the budget share for food and income. However, this correlation is not perfect. For example, comparing households one and two, per capita income is the same while budget share for food is lower for the second household. Comparing household three with four and five with six, per capita income and budget share for food both change in the same direction. Thus although income is an important factor, it is clear that demographic variables are also important for explaining the budget share pattern for food.

Budget share patterns for disaggregated foods over the household profiles predicted by the model are shown in Figure 2. Food away from home is the largest budget item for all households, and thus in a sense is the driving force behind changes in the budget shares of the other food products since budget shares sum to one. The budget share for food away from home is largest for the single young male and low for the married couple with two young children and only one working spouse, the household with both partners retired, and the household with the widowed spouse.

The budget share for dairy products predictably is greater when there are children in the household. As noted previously, the income elasticity for dairy products is quite small, so the differences observed among households is due to changes in the demographic composition. It is also interesting to note that the income and household size for households one and nine (single young male and elderly widow) are identical, but that there is a marked difference in the budget share pattern. This further illustrates the importance of household demographic composition.

Figure 2 Disaggregated food budget shares for profile households



Conclusions

Much of the literature on incorporating demographic variables into demand models focuses on modifying the form of the utility function. This approach is, however, relatively restrictive, and as a result, the direct modification of a demand system offers greater flexibility for empirical research. The standard procedure is to define a household equivalence scale that affects the parameters of the demand system, but the lack of a sound theoretical basis for household decision-making implies that there is little guidance from theory on how this step should be done.

Tyrrell's specification of a continuous equivalence scale is used for this research to avoid the difficulty of choosing boundary values for age groups, and to allow for interactions between household size and the distribution of ages of individual household members. In addition, Tyrrell's specification is extended to include a broad range of demographic characteristics such as race, life style and employment status.

Demand systems for nine food groups are fitted using the DM model and data from the I986 CEDS. Different models are fitted for eight season/ region combinations, and prices are assumed to be homogeneous within each of these groups. The unrestricted specification performs well in terms of goodness of fit. Likelihood ratio tests for simpler representations of the demographic structure of the model are rejected. Testing also rejects simplification of the season/region structure of the model, implying that differences among seasons and regions cannot be attributed solely to price differences (i.e. different intercepts).

An analysis of budget share patterns predicted by the model for nine profile households at various stages of the household life-cycle shows how sensitive household expenditures on food are to changes in demographic variables. This is particularly important for food away from home.

REFERENCES

Barnes, Roberta, and Gillingham, Robert. "Demographic Effects in Demand Analysis: Estimation of the Quadratic Expenditure System Using Microdata." The Review of Economics and Statistics 66 (November, 1984): 591-601.

Barten, A.P. "Consumer Demand Functions Under Conditions of Almost Additive Preferences." Econometrica 32 (1964): 1-38.

Blaylock, James R., Smallwood, David M. U.S. Demand for Food: Household Expenditures, Demographics, and Projections. United States Department of Agriculture, Economic Research Service, Technical Bulletin No. 1713, Washington, D.C. 20005-4788, February 1986.

Blokland, J. Continuous Consumer Equivalence Scales: Item-specific Effects of Age and Sex of Household Members in the Budget Allocation Model. The Hague: Martinus Nijhoff, 1976.

Buse, Rueben; Price, David W.; West, Donald A.; Price, Dorothy Z.; Chavas, Jean-Paul; Blaylock, James; and Smallwood, David. "Socioeconomic, Demographic, and Psychological Variables in Demand Analyses." pp. 186-215. In Food Demand Analysis: Problems, Issues and Empirical Evidence. Edited by Raunikar, Robert and Huang, Chung-Liang. Ames: Iowa State University Press, 1987.

Buse, Rueben C., and Salathe, Larry E. "Adult Equivalent Scales: An Alternative Approach." American Journal of Agricultural Economics 60 (August, 1978): 460-468.

Deaton, Angus, and Muellbauer, John. Economics and Consumer Behavior Cambridge: Cambridge University Press, 1980a.

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

Engel, E. "Die Lebenskoten Belgischer Arbeiter-Familien fruher and jetzt." International Statistical Institute Bulletin 9 (1895): 1-74.

Gorman, W.M., "Some Engel Curves." In Essays in the Theory and Measurement of Consumer Behavior. Edited by A.S. Deaton. New York: Cambridge University Press, 1981.

Gorman, W.M. "Tricks with Utility Functions." pp. 211-243. In M.J. Artis and A.R. Nobay, eds., Essays in Economic Analysis. The Proceedings of the Association of University Teachers of Economics Sheffield 1975. New York: Cambridge University Press, 1976.

Gould, Brian, Cox, Thomas and Perali, Federico. "Demand for Food Fats and Oils: The Role of Demographic Variables and Government Donations." American Journal of Agricultural Economics 73 (February 1991): 212-221.

Harvey, A.C. The Econometric Analysis of Time Series Oxford: Philip Allan Publishers Ltd., 1981

Heien, Dale, and Pompelli, Greg. "The Demand for Beef Products: Cross-Section Estimation of Demographic and Economic Effects." Western Journal of Agricultural Economics 13 (1988): 37-44.

Kokoski, Mary. "An Empirical Analysis of Intertemporal and Demographic Variations in Consumer Preferences." American Journal of Agricultural Economics 68 (November, 1986): 894-907.

Lewbel, Arthur. "Additive Separability and Equivalent Scales." Econometrica 54 (January, 1986): 219-222.

Lewbel, Arthur. "A Unified Approach to Incorporating Demographic or Other effects into Demand Systems." Review of Economic Studies 52 (1985): 1-18.

McCracken, Vicki A. and Brandt, Jon A. "Household Consumption of Food-Away-From-Home: Total Expenditure and by Type of Food Facility." American Journal of Agricultural Economics 69 (May, 1987): 274-284.

Muellbauer,J. "The Estimation of the Prais-Houthakker Model of Equivalence Scales." Econometrica 48 (January, 1980): 153-176.

Muellbauer, J. "Testing the Barten Model of Household Consumption Effects and the Cost of Children." Economic Journal 87 (September, 1977): 460-487.

Muellbauer, J. "Household Composition, Engel Curves, and Welfare Comparisons Between Households: A Quality Approach." European Economic Review 5 (1974): 103-122.

Nelson,Julie. "Household Economies of Scale in Consumption: Theory and Evidence." United States Bureau of Labor Statistics Working Paper No. 171. Washington,D.C. 20212, August 1987 (Rev. December 1987).

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

Pollak, R.A., and Wales, T.J. "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects," Econometrica 8 (1980): 575-612.

Pollak, R.A., and Wales, T.J. "Estimation of Complete Demand Systems from Household Budget Data: The Linear and Quadratic Expenditure Systems." American Economic Review 68 (1978): 348-359.

Prais, S.J., and Houthakker, H.S. The Analysis of Family Budgets Cambridge: Cambridge University Press (1955); 2d. ed., 1971.

Price, D. W. "The Effects of Household Size and Composition on the Demand for Food." In Food Demand Analysis: Implications for Future Consumption. Edited by Oral Capps, Jr. and Benjamin Senauer. Virginia Polytechnic Institute and State University, Blacksburg, Virginia, August 1986.

Raunikar, Robert and Huang, Chung-Liang, Editors. "Socioeconomic, Demographic, and Psychological Variables in Demand Analyses." In Food Demand Analysis: Problems, Issues, and Empirical Evidence, pp. 186-215. Ames: Iowa State University Press, 1987.

Schrimper, Ronald. "Effects of Increasing Elderly Population on Future Food Demand and Consumption." In Food Demand Analysis: Implications for Future Consumption Edited by Oral Capps Jr. and Benjamin Senauer. Virginia Polytechnic Institute and State University, August 1986.

Tedford, John R.; Capps, Oral, Jr.; and Havlicek, Joseph, Jr. "Adult Equivalent Scales Once More -- A Developmental Approach." American Journal of Agricultural Economics 68 (May 1986): 322-333.

Tedford, John R.; Capps, Oral, Jr.; and Havlicek, Joseph, Jr. "Regional Equivalence Scales for Convenience Foods." Journal of the Northeastern Agricultural Economics Council 13 (April 1984): 33-39.

Tyrrell, T.J. "An Application of the Multinomial Logit Model to Predicting the Pattern of Food and Other Household Expenditures in the Northeastern United States." Ph.D. dissertation, Cornell University, 1979.

Tyrrell, T.J, and Mount, T.D. "Analysis of Food and Other Expenditures Using a Linear Logit Model." In Food Demand Analysis: Problems, Issues, and Empirical Evidence, pp. 143-53. Edited by Robert Raunikar and Chung-Liang Huang. Ames: Iowa University Press, 1987.

Woodland, A.D. "Stochastic Specification and the Estimation of Share Equations." Journal of Econometrics 10 (1979): 361-383.

OTHER AGRICULTURAL ECONOMICS WORKING PAPERS

No. 92-01	A Theory of Competition Between Politicians for Political Support	Jo Swinnen Harry de Gorter
No. 92-02	Negotiating a Response to Climate Change: The Role of Biological Emissions	Thomas E. Drennen Duane Chapman
No. 92-03	Environment, Income, and Development in Southern Africa	Duane Chapman
No. 92-04	Data Requirements for Energy Forecasting	Timothy D. Mount
No. 92-05	Oil Shocks and the Demand for Electricity	Edward C. Kokkelenberg Timothy D. Mount
No. 92-06	The Effects of Carbon Taxes on Economic Growth and Structure Will Developing Countries Join a Global Agreement?	Steven C. Kyle
No. 92-07	Climate Change, Sustainable Economic Systems and Welfare	Timothy D. Mount
No. 92-08	Welfare Effects of Improving End-Use Efficiency: Theory and Application to Residential Electricity	Jesus C. Dumagan Timothy D. Mount
No. 92-09	The Value of Segmenting the Milk Market into bST-Produced and Non-bST-Produced Milk	Loren W. Tauer
No. 92-10	World Oil: Hotelling Depletion, or Accelerating Use?	Duane Chapman