QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 60-69

THE IMPORTANCE OF DEMOGRAPHIC VARIABLES IN THE MODELING OF FOOD DEMAND

Hanna Dudek
Department of Econometrics and Statistics, Warsaw University of Live Sciences
e-mail: hanna_dudek@sggw.pl

Abstract: The general objective of this research is to assess the impact of demographic variables on food demand in Poland. The empirical analysis of this paper is based on the household data, collected by GUS (Central Statistical Office) in the years 2001-2004 (household budget data).

Keywords: complete demand system, demographic variables, elasticity of demand

INTRODUCTION

During the last three decades, consumer demand analysis has moved toward system-wide approaches. Increasing attention has been given to the estimation of complete demand systems that consistently account for the interdependence in the choices made by consumers between a large number of commodities. Many algebraic specifications of demand systems have been developed, including the linear and quadratic expenditure systems, the Rotterdam model, Translog models and the Almost Ideal Demand System (AIDS) or its quadratic extension (QUAIDS) [Barnett, Serletis 2008].

The objective of this study is to estimate the impact of economic factors, such as the prices and the expenditures, and noneconomic factors, i.e. the demographic variables, on household demand for eight aggregated food items in Poland.

The issue analysed in this paper is particularly important in the case of the situation in Poland, where the **single-equation** approach to demand analysis dominates. The use of single-equation models suffers from some shortcomings, first of all it ignores the cross-equation restrictions implied by the neoclassical theory of consumer behavior. Following paper presents a struggling attempt to fill the gap in the literature in regard to system-wide approach to food expenditures in

Poland. Moreover, it emphasizes the incorporation of demographic variables in the analysis of food demand, which is very important due to changes demographic profile of consumers, i.e. the problem of aging society. As far as the author is concerned such research has not been conducted for the Polish household data.

ECONOMETRIC DEMAND ANALYSIS

The articles written by H. Working in the 1940s and C. Leser in the 1960s were one of the first papers on econometric demand analysis. Nonetheless, their significant contribution to the demand analysis was not consistent with the utility maximization theory. According to this theory, a sample of households behaves as a representative consumer who maximizes his utility function $u(\mathbf{q})$ subject to the budget constraint $\mathbf{p'q} = x$ where:

$\mathbf{q}$ is the vector of food demanded, $\mathbf{q}=[q_1, q_2,..., q_n]$,

$\mathbf{p}$ is the corresponding vector of prices, $\mathbf{p}=[p_1, p_2,..., p_n]$,

x is the total expenditures to consume $\mathbf{q}$.

By solving this maximization problem, we obtain a system of $n+1$ demand equations specified as follows (see e.g. [Varian 1992], [Barnett, Serletis 2008]):

$$\mathbf{q} = \mathbf{q}(\mathbf{p}, x) \tag{1}$$

The solution to the utility maximization problem yields a set of ordinary demand curves conditional on given prices and income. The system (1) is assumed to satisfy the theoretical plausibility conditions, especially adding up, homogeneity, symmetry and negativity[1] [Edgerton et al. 1996].

Only coherent demand systems allow to model various consumption patterns and behavior sufficiently, while simultaneously satisfying restrictions given by the economic theory. Such examples of coherent demand system present the Linear Expenditure System (LES) [Stone 1954], Transcendental Logarithmic System [Jorgenson, Lau Stoker 1982], Almost Ideal Demand System and its quadratic extension [Banks et al. 1997].

In order to illustrate the incorporation of the demographic variables the AIDS model (Almost Ideal Demand System) is used in this paper. It would also be reasonable to consider other demand systems such as QUAIDS (Quadratic Almost Ideal Demand System) and QES (Quadratic Expenditure System), however,

[1] These conditions represent the basic restrictions imposed on all demand functions $q_i = q_i(\mathbf{p}, x)$, $i = 1, 2, .., n$:

- the **adding up restriction** implies that the budget x is totally used;
- the homogeneity condition requires the demand functions to be homogeneous of degree zero in both prices and total expenditures;
- the symmetry and negativity **restrictions** imply that the substitution matrix should be symmetric and negative semidefinite.

estimation difficulties relating to the implementation of nonlinear numerical procedures, which are characteristic to these models, cause that no such attempt has been undertaken[2]. Though, it appears that in order to achieve the objective of this paper, it is sufficient to use a simpler model such as AIDS.

Apart from its flexibility, the main advantages of the AIDS model are as follows: first, it allows an exact aggregation among consumers; second, there is a possibility to estimate a non-linear model; third, it is a popular model because of its empirical validation. The general specification of the AIDS model is given by [Deaton, Muellbauer 1980]:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(x/P) \tag{2}$$

where:

w_i is the expenditure share associated with the ith good, $i=1, 2,\dots,n$,

α_i is the constant coefficient in the ith share equation,

γ_{ij} is the slope coefficient associated with the jth good in the ith share equation,

p_j is the price of the jth good,

x is the total expenditure on the system of goods given by the following equation:

$x = \sum_{i=1}^{n} p_i q_i$, where q_i is the quantity demanded for the ith good,

P is the general price index defined by:

$$\log P = \alpha_0 + \sum_{i=1}^{n} \alpha_i \log p_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \log p_i \log p_j \tag{3}$$

In empirical studies, in order to avoid the non-linearity and reduce the multi-colinearity effects in the model, the equation (3) is often approximated by a Stone index defined as [Deaton, Muellbauer 1980]:

$$\log P^* = \sum_{i=1}^{n} w_i \log p_i \tag{4}$$

Model AIDS (2) with the price index (4) instead (3) is called LA/AIDS model[3] (Linear Approximation of the Almost Ideal Demand System).

[2] Considering that eight expenditure groups were taken into account in the paper, the estimation of nonlinear models using an ordinary personal compute would be extremely time-consuming.

[3] The Stone index is one of a numerous of indices that could be used to define a LA/AIDS specification. The discussion relating to disadvantages of Stone index can be found in [Moschini 1995], [Asche, Wessells 1997] and [Barnett, Seck 2008]. Despite various shortcomings, this index is still applied in empirical researches (see e.g. [Suchecki 2006], [Regorsek, Erjavec 2007].

In our empirical estimation the usual theoretical restrictions derived from the utility maximization and the demand theory are directly imposed into LA/AIDS parameters. These restrictions are:

$$\sum_{i=1}^{n}\alpha_i = 1, \sum_{i=1}^{n}\beta_i = 0, \sum_{i=1}^{n}\gamma_{ij} = 0 \text{ for adding up,} \tag{5}$$

$$\sum_{j=1}^{n}\gamma_{ij} = 0 \text{ for homogeneity and,} \tag{6}$$

$$\gamma_{ij} = \gamma_{ji} \text{ for symmetry.} \tag{7}$$

According to Green and Alston (1990), elasticities in LA/AIDS can be expressed as: $e_i = 1 + \frac{\beta_i}{w_i}$ for income elasticity and $e_{ij}^{*} = -\delta_{ij} + w_j + \gamma_{ij} / w_i$ for compensated[4] (Hicksian) elasticity. The uncompensated (Marshallian) elasticity of relative expenditures on the commodity i relative to the price of commodity j is given as: $e_{ij} = -\delta_{ij} + \frac{\gamma_{ij}}{w_i} - \beta_i \frac{w_j}{w_i}$, where δ_{ij} is the Kronecker delta taking the value 1 for $i=j$ and 0 for $i \neq j$ [Dudek 2008].

DEMOGRAPHIC VARIABLES

There are general procedures for taking into account the demographic variables into classes of demand system [Pollack, Wales 1981; Ray 1983]. If we denote the originally demand system by $w_i = f(\mathbf{p}, logx)$, which depends on the vector of prices $\mathbf{p}$ and on the total expenditures x, then:

1) the method known as "demographic scaling" transforms it into $w_i = m_i \cdot f(\mathbf{p}, logx)$, where the m's are scaling parameters, which depend on the demographic variables [Barten 1964];

2) the procedure named "demographic translating" replaces and translates the original demand system $w_i = f(\mathbf{p}, logx)$ into $w_i = d_i + f(\mathbf{p}, logx)$, where d is a parameter depending on the vector of demographic variables [Pollak, Wales 1979];

[4] The economic theory distinguishes two types of demand and thus the elasticities. *Compensated (or Hicksian) elasticity* derived from Hicks (compensated) demand, measures only the substitution effect, i.e. the change in demand due to the change in relative prices if the effect on real income due to the change in prices is compensated. *Marshallian demand and elasticity* considers, however, not only the substitution effect caused by the relative prices change, but also the income effect arising from the change in real income due to the change in prices.

3) the Gorman's specification proposed a method that replaces the original demand system in the following: $w_i = d_i + m_i \cdot f(\mathbf{p}, logx)$, where the d's and the m's depend on the demographic variables [Gorman 1976];
4) the "reverse Gorman" specification can be obtained by firstly by translating, then scaling to yield the following demand system: $w_i = m_i \cdot (d_i + f(\mathbf{p}, logx)$ [Pollak, Wales 1981];
5) the "price scaling" technique replaces the original demand system by $w_i = f\left(\mathbf{p}, \frac{\log x}{m(\mathbf{p},\mathbf{z})}\right) + \frac{\partial \log(m(\mathbf{p},\mathbf{z}))}{\partial \log p_i}$, where $m(\mathbf{p},\mathbf{z})$ - the scaling factor, $\mathbf{z}$ - the vector of demographic variables [Ray 1983].

The procedures mentioned above are general in the sense that they do not assume a particular form of the original demand system, but can be used in conjunction with any complete demand system. Arranging these procedures in an unambiguous ranking is not possible [Pollak, Wales 1981]. First of all, some of them are not nested. Furthermore, their assessment depends also on the functional form used to estimate the demand system. Note that the estimation of some procedures is computationally complex, thus only one relatively simple procedure, "demographic translating", was taken into account in this paper.

LA/AIDS models, as originally proposed by Deaton and Muellbauer (1980), did not consider the demographic variables. However, such variables appear to be crucial in household survey data, in which economic responses to the price changes can be considerably influenced by all the sorts of personal or household effects. The study takes into consideration the "demographic translating" in the LA/AIDS model by allowing the intercept term of each equation to be a function of the demographic variables:

$$w_i = d_i + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(x / P^*) \tag{8}$$

where $d_i = \alpha_{0i} + \sum_{k=1}^{K} \alpha_{ki} z_k$, K – a number of demographic variables.

The term demographic variables here denotes:

$$z_j = \frac{m_j}{m}, j=1,..,K-1, \qquad z_K = \ln(m) \tag{9}$$

where m_j – a number of household members in age group j,
m – a household size (number of household members).

DATA AND ESTIMATION

The estimation is based on a sample of the household data, derived from the Polish Household Budget Survey5 conducted by the Central Statistical Office of Poland annually conducted by the Central Statistical Office of Poland. The data mostly come from the survey on household monthly expenditures for the years 2001-2004, latest data is not available. Only the set of monthly prices indices was taken from publication *Prices in the National Economy in 2001-2004*.

Generally, data derived from National Household Budget Surveys often causes so-called problem of "zero expenditures" due to the fact that not every single household adhering to the sample buy at least one commodity from each of the aggregated groups. The reasons for this phenomenon are following: the infrequency of the purchase, the seasonality of some products, the self-production of some commodities, etc. Accordingly, data derived from the household of employees[6], in which the problem of "zero expenditures" was not so significant as in remaining groups, are only taken into account.

The specific food commodities within the food groups used in the empirical analysis are: 1) bread and cereals, 2) meat and fish, 3) milk, cheese and eggs, 4) oils and fats, 5) fruit, 6) vegetables, 7) sugar, jam, honey, chocolate and confectionery, 8) other food products. For such groups the ratio of households with "zero expenditures" did not exceed 3%, thus the problem of the censored data was passed over in this study.

The LA/AIDS model was estimated by the seemingly unrelated regression (SUR) technique in the STATA version 10 statistical package. The homogeneity and symmetry restrictions were imposed on the estimated model. To avoid singularity derived from adding-up constraint in the variance-covariance matrix one equation was deleted from direct estimation in the demand system. The parameters' estimates of this equation were recovered using homogeneity, symmetry and adding-up conditions.

RESULTS

We considered demand systems with various set of the demographic variables. In accordance with the statistical criteria, **i.e.** statistical significance of parameters, **Akaike** and **Schwarz** information criteria, one model with two

---

[5] The Household Budget Surveys data are considered superior compared to the available time-series data for the research because they include detailed demographic characteristics that allow heterogeneity in preferences across households. Additionally, the large sample size included in the NBS household survey data allows estimating a relatively large demand system.

[6] The number of households of employees participating in the survey in each year was about 12500.

demographic variables, **i.e** ratio number of household members aged 14 years and over[7] and household size, was chosen. Most of the parameter estimates were significant at the 0,05 significance level. For reasons of space, all detailed results could not be presented. The complete results are available from the author upon request. The obtained results[8] concerning the parameter estimates, standard errors and the goodness of fit measures are presented in Table 1.

Table 1. Estimation result of LA/AIDS model

Variable	Estimates of parameters in equation:							
	1	2	3	4	5	6	7	8
z_1	-0,0575*	0,0760*	-0,0371*	0,0066*	0,0165*	0,0223*	-0,0303*	0,0033*
	(0,0014)	(0,0021)	(0,0013)	(0,0006)	(0,0009)	(0,0011)	(0,0009)	(0,0006)
z_2	0,0018*	-0,0001	-0,0003	-0,0003	-0,0005*	-0,0019*	0,0010*	0,0003
	(0,0006)	(0,0001)	(0,0003)	(0,0002)	(0,0002)	(0,0005)	(0,0004)	(0,0002)
log(x/P)	-0,0310*	0,0416*	-0,0135*	-0,0059*	0,0046*	0,0051*	0,0008*	0,0684*
	(0,0006)	(0,0009)	(0,0006)	(0,0002)	(0,0004)	(0,0005)	(0,0004)	(0,0005)
$\log p_1$	0,1083*	-0,0653*	0,2459*	-0,2173*	0,0083*	-0,0359*	0,0130*	-0,0569*
	(0,0351)	(0,0094)	(0,0179)	(0,0100)	(0,0036)	(0,0021)	(0,0086)	(0,0161)
$\log p_2$	-0,0653*	0,5870*	-0,7930*	0,4655*	-0,0193*	0,0541*	-0,2102*	-0,0188*
	(0,0094)	(0,0114)	(0,0083)	(0,0043)	(0,0036)	(0,0025)	(0,0050)	(0,0042))
$\log p_3$	0,2459*	-0,7930*	0,8180*	-0,5008*	-0,0686*	-0,0462*	0,3024*	0,0423*
	(0,0179)	(0,0083)	(0,0160)	(0,0071)	(0,0031)	(0,0019)	(0,0066)	(0,0082)
$\log p_4$	-0,2173*	0,4655*	-0,5008*	0,4775*	0,0578*	0,0121*	-0,3176*	0,0227*
	(0,0054)	(0,0043)	(0,0071)	(0,0060)	(0,0015)	(0,0008)	(0,0042)	(0,0049)
$\log p_5$	0,0083*	-0,0193*	-0,0686*	0,0578*	[0,0115*	0,0370*	0,0057*	-0,0093*
	(0,0036)	(0,0036)	(0,0031)	(0,0015)	(0,0022)	(0,0011)	(0,0020)	(0,0015)
$\log p_6$	-0,0359*	0,0541*	-0,0462*	0,0121	0,0370*	0,0053*	-0,0263*	0,0001
	(0,0021)	(0,0025)	(0,0019)	(0,0008)	(0,0011)	(0,0014)	(0,0012)	(0,0002)
$\log p_7$	0,0130	-0,2102*	0,3024*	-0,3176*	0,0057*	-0,0263*	0,2414*	-0,0083*
	(0,0086)	(0,0050)	(0,0066)	(0,0042)	(0,0020)	(0,0012)	(0,0048)	(0,0043)
$\log p_8$	-0,0569*	-0,0188*	0,0423*	0,0227*	-0,0093*	0,0001	-0,0083*	0,0283*
	(0,0161)	(0,0042)	(0,0082)	(0,0049)	(0,0015)	(0,0002)	(0,0043)	(0,0088)
const	0,4288*	0,0100	0,2804*	0,0861*	0,0133*	0,0519*	0,0842*	0,0453*
	(0,0040)	(0,0061)	(0,0038)	(0,0017)	(0,0027)	(0,0034)	(0,0027)	(0,0017)
	Goodness of fit							
R^2	0,0680	0,0736	0,1636	0,1865	0,0038	0,0269	0,0139	0,0829

Source: the author's own computations in the STATA statistical package; z_1 denotes ratio number of adults, z_2 – logarithm of household size; standard error in parentheses; asterisk indicates significance at 0,05; food products are following: 1) bread and cereals, 2) meat and fish, 3) milk, cheese and eggs, 4) oils and fats, 5) fruit, 6) vegetables, 7) sugar, jam, honey, chocolate and confectionery, 8) other food products

[7] This is consistent with the OECD equivalence scale in which members of household aged less than 14 are considered children and members aged 14 and over - adults.

[8] STATA's *sureg* command was used for the LA/AIDS estimation.

The log likelihood ratio test[9] and the t-test show that the inclusion of a demographic variables was justified. The implication is that various types of a household, the different composition and the age structure, have an impact on food demand. Table 2 presents results of total food expenditure elasticities[10].

Table 2. Estimated food expenditure elasticities for chosen types of households[11]

	One adult without children	One adult + one member aged 14 and over Number of children (aged below 14) 0	1	2
bread and cereals	0,8508 [0,8442;0,8574]	0,8346 [0,8273;0,8419]	0,8669 [0,8610;0,8728]	0,8586 [0,8529;0,8648]
meat and fish	1,1483 [1,1416;1,1549]	1,1283 [1,1226;1,1341]	1,1583 [1,1512;1,1654]	1,1406 [1,1343;1,1469]
milk, cheese and eggs	0,9261 [0,9190;0,9331]	0,9189 [0,9111;0,9266]	0,9313 [0,9247;0,9378]	0,9279 [0,9211;0,9348]
oils and fats	0,8997 [0,8902;0,9093]	0,9018 [0,8924;0,9111]	0,8884 [0,8778;0,8991]	0,8934 [0,8837;0,9041]
fruit	1,0508 [1,0390;1,0625]	1,0567 [1,0436;1,0699]	1,0636 [1,0489;1,0783]	1,0690 [1,0530;1,0849]
vegetables	1,0317 [1,0217;1,0417]	1,0306 [1,0210;1,0402]	1,0323 [1,0222;1,0425]	1,0357 [1,0245;1,0470]
sugar, jam, honey, chocolate and confectionery	1,0011 [0,9889;1,0134]	1,0013 [0,9874;1,0402]	1,0010 [0,9904;1,0116]	1,0011 [0,9898;1,0123]
other food products	0,8976 [0,8841;0,9111]	0,8952 [0,8814;0,9090]	0,8784 [0,8623;0,8944]	0,8885 [0,8738;0,9032]

Source: the author's own computations in the STATA statistical package. The values included **in parentheses** are the **confidential intervals** (95%) of the food expenditure elasticities

The results presented in Table 2 reveal that all the food groups are fairly sensitive to the food expenditure changes. Moreover, the elasticities can vary according to the different demographic profiles that exist in the population[12].

The study has found some differences in the elasticity estimates for various demographic types of household. For example, the elasticity of meat and fish for households with two adults significantly differs from other considered types –

[9] The null hypothesis of overall absence of demographic effects is strongly rejected – LR=4433 and critical value $\chi^2(0,05;14) = 24$.

[10] For reasons of space, results of the price elasticities are not presented.

[11] All elasticities are evaluated at group's means.

[12] It should be noted that the demand for food in Poland, apart from demographic attributes of households, probably depends on other numerous relevant attributes of households such as, for example place of living, educational attainments and occupation of his members. The impact of such features on food demand in Poland can form a subject of research in the future.

households consisted of: a single person and two adults with one child or two children.

CONCLUDING REMARKS

The demographic variables played a major role in the analysis of the household budget data. Instead of assuming that all the households in the sample have identical tastes, only those with the same demographic profiles are assumed to have the same demand functions. The household size and its composition have been used as the demographic variables in demand studies, although seldom in the context of complete system of demand equations. In this study, we used the "demographic translation" procedure to incorporate demographic variables into the demand system. This procedure allows to explain the heterogeneous nature in the household consumption patterns.

Eventually, the results provide more insight for the understanding of the household consumption habits in Poland. It can be valuable for marketing strategies because a strong segmentation among households provides a more comprehensive picture of the food expenditures. In conclusion, this study indicates that the changing demographic profile of consumers in Poland has had a significant impact on food demand.

REFERENCES

Asche F., Wessells C. R. (1997) On Price Indices in the Almost Ideal Demand System, American Journal of Agricultural Economics, vol. 74, pp. 1182-1184.

Banks J., Blundell R., Lewbel A. (1997) Quadratic Engel Curves and Consumer Demand, The Review of Economics and Statistics, Vol. 74, pp. 527-539.

Barnett W. A., Seck O. (2008) Rotterdam Model versus Almost Ideal Demand System: Will the Best Specification Please Stand up?, Journal of Applied Econometrics, Vol. 23, pp. 699-728.

Barnett W. A., Serletis A. (2008) Consumer Preferences and Demand Systems, Journal of Econometrics, Vol. 147, No. 2, pp. 210-224.

Barten A. P. (1964) Family Composition, Prices and Expenditure Patterns, in: P.E. Hart, F.Mills, J.K. Whitaker (eds.), Econometric Analysis for National Economic Planning, Butterworths, London.

Deaton A. Muellbauer J. (1980) An Almost Ideal Demand System, American Economic Review, Vol. 70, No 3, pp. 312-316.

Dudek H. (2008) Price Elasticities of Food Demand - Analysis of LA/AIDS Model, Annals of the Polish Association of Agricultural and Agribusiness Economists, Vol. 10, No. 4, pp. 62-67 (in Polish).

Edgerton D. L., Assarsson B., Hummelmose A., Laurila I P., Rickertsen K., Vale P. H. (1996) The Econometrics of Demand Systems - With Applications to Food Demand in the Nordic Countries, Kluwer Academic Publisher, Dordrecht-Boston-London.

Gorman W. (1976) Tricks with Utility Functions, in: M. Artis, A. Nobay (eds.), Essays in Economic Analysis, Cambridge University Press, Cambridge.

Green R. Alston J. M. (1990) Elasticities in AIDS Models, American Journal of Agricultural Economics, Vol. 72, pp. 442-445.

Jorgenson D. W., Lau L. J., Stoker T. M. (1982) The Transcedental Logarithmic Model of Aggregate Consumer Behavior, in: R. Baumann, G. Rhodes, (eds.), Advances in Econometrics, JAI Press, Vol. 1, pp. 97-238.

Leser C. E. (1963) Forms of Engel Functions, Econometrica, Vol. 31, No. 4, pp. 694-703.

Moschini G. (1995) Units of Measurement and the Stone Index in Demand System, American Journal of Agricultural Economics, Vol. 77, No. 1, 63-68.

Pollak R. A., Wales T. J. (1979) Welfare Comparison and Equivalence Scales, American Economic Review, Vol. 69, pp. 216-221.

Pollak R. A., Wales T. J. (1981) Demographic Variables in Demand Analysis, Econometrica, Vol. 49, No. 6, pp. 1533-1551.

Ray R. (1983) Measuring the Cost of Children: an Alternative Approach, Journal of Public Economics, Vol. 22, pp. 89-102.

Regorsek D., Erjavec E. (2007) Food Demand in Slovenia, Acta Agriculturae Slovenica, Vol. 89, No. 1, pp. 221-232.

Stone J. R. N. (1954) Linear Expenditure Systems and Demand Analysis: an Application to the Pattern of British Demand, Economic Journal, Vol. LXIV, No. 255, pp. 511-527.

Suchecki B. (2006) Complete Demand Systems, Polskie Wydawnictwo Ekonomiczne, Warszawa (in Polish).

Varian H. R. (1992) Microeconomic Analysis, W.W. Norton&Company, New York.

Working H. (1943) Statistical Laws of Family Expenditure, Journal of the American Statistical Association, Vol. 38, pp. 43-56.