JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **15**: 275–287 (2000)

LINEAR HOUSEHOLD TECHNOLOGIES

C. ANDREA BOLLINO[a], FEDERICO PERALI[b]* AND NICOLA ROSSI[c]
[a] *Luiss, Viale Pola 12, 00198 Rome, Italy*
[b] *Department of Economics, University of Verona, Via dell'Artigliere 19, 37129 Verona, Italy*
[c] *Department of Economic Studies, University of Rome 'Tor Vergata', 00133 Rome, Italy*

SUMMARY

This paper discusses, estimates and formally compares the best known procedures for incorporating demographic variables into complete demand systems. In particular, a class of general procedures belonging to Gorman's family of 'general linear household technologies' is introduced. Estimation and comparison of different procedures make use of Italian household budget data for the years 1973–1992, incorporating a single demographic variable (family size) into a Generalized Quadratic Almost Ideal Demand System. In our empirical example, however, even the most general household technologies are unable to fully capture the behavioural heterogeneity shown by the data. Copyright © 2000 John Wiley & Sons, Ltd.

1. INTRODUCTION

Conditioning variables such as age, number and age composition of children, education, social status, residential location, etc. play a major role in determining household consumption patterns. Indeed, substantial efforts have been devoted by demand analysts to the 'explanation' of differences in tastes in terms of observable differences in household characteristics. Muellbauer (1977, 1980), Pollak and Wales (1978, 1980, 1981), and Ray (1985, 1993, 1996) have described and estimated general procedures for incorporating demographic variables into demand systems within a utility consistent framework. Their work, however, (based, in all cases, on the UK Family Expenditure Survey) focuses on only some of the alternatives suggested by the theoretical literature. Furthermore, it does not always provide an unambiguous ranking of procedures.

A common view holds that following such a structural approach when incorporating demographics into a demand system may be a useless research strategy. After all, if a household production approach cannot be distinguished from a situation where preferences depend on demographics, why not simply let each parameter be a linear or non-linear function of the demographic variables? Or stratify and estimate different systems for more or less broadly defined socio-demographic groups?

Apart from efficiency implications, this line of reasoning overlooks the fact that 'simply' letting parameters vary with demographics may have substantial behavioural implications: for example, a demographically changing constant term in an Almost Ideal demand system entails a rather different notion of fixed costs with respect to a demographically changing constant term in a Generalized Almost Ideal demand system. In turn, different notions of fixed costs may possess rather different implications in terms of applied welfare analysis.[1]

* Correspondence to: F. Perali, Department of Economics, University of Verona, Via dell'Artigliere 19, 37129 Verona, Italy; e-mail: wicaro@chiostro.univr.it

[1] For example, some demographic procedures may be understood as a complete affine transformation of the original demand system implying cardinal full comparability. Others may not.

Received 5 July 1996
Revised 15 July 1999

Therefore, notwithstanding the above criticism, this paper takes yet another look at the structural approach to incorporate demographics into demand systems. Following the path set out by Muellbauer, Pollak, Wales and Ray, it discusses, estimates, and formally compares most known general procedures for incorporating demographic variables into complete demand systems, where, according to Pollak and Wales (1981, p. 1533), a 'general procedure' is one that does not assume the original demand system to have a particular functional form.[2]

In doing so, a class of general procedures, belonging to Gorman's family of 'general linear household technologies' and never considered before in the published literature is discussed and estimated. These procedures allow for joint demographically varying effects across commodities. They nest the best-known procedures and compare then favourably in an economic as well as in a statistical sense.

To estimate and compare the different procedures we make use of the published Italian Household Budget Survey for the period 1973–1992, incorporating a single demographic variable (the number of household members)[3] into a Generalized Quadratic Almost Ideal demand system.

While addressing the question of the specification of which parameters depend on the demographic variables and which do not, the paper also investigates the related question of the specification of the functional form relating demographically varying parameters to demographic variables. Indeed, the two issues can hardly be distinguished: depending on the underlying procedure, linear demographic functions can lead to both linear and non-linear demographic effects. It will be suggested that, on the basis of the available evidence, non-linear demographic effects arising from non-linear demographic functions tend to dominate non-linear demographic effects based on linear demographic functions.

In short, the results show that more demographic structure is better than merely some structure. However, whatever their degree of generality, household technologies — based on a single demographic input such as family size in our case — will be shown to be apparently unable to capture the vast amount of behavioural heterogeneity present in the data.

2. GENERAL PROCEDURES

In general terms, the introduction of demographic effects into complete demand systems can be described along the lines originally suggested by Gorman (1976) and formalized by Lewbel (1985), i.e. in terms of household production technologies. In this context, households with different socio-demographic characteristics are assumed to share the same stable structure of references defined over a given number of intermediate goods not available on the market (such as prepared meals, transportation, and so on). Households differ, however, in the (demographically varying) production technology that enables them to construct intermediate goods from input goods available on the market. In other words, the focus of the analysis shifts from

[2] For example, the Prais and Houthakker (1955) procedure is not a 'general' one in the sense previously stated (see Muellbauer, 1980; Pollak and Wales, 1981; Lewbel, 1986). In addition, notice also that the Rothbarth approach is not fully general, insofar as it imposes a specific separability structure between goods consumed by different types of household (Browning, 1992; Nelson, 1992; Rothbarth, 1943; Tsakloglu, 1991).

[3] Admittedly, family size is one of the crudest ways of describing household composition. Age and number of children would certainly be a better and more appropriate alternative especially as far as the issue of economies of scale in consumption is concerned. Unfortunately, the Italian Central Statistical Office does not supply the latter information for the whole period under consideration. It is interesting to note that this is mainly because some systems of social benefits, like the Italian one, still focus explicitly on family size.

differences in preferences to differences in technological constraints faced by households with different socio-demographics profiles.

Let $q = (q_1, \ldots, q_I)$ be the quantity (column) vector of input goods bought by consuming units and $p = (p_1, \ldots, p_I)$ the corresponding price vector. Then $y = p'q$ denotes total expenditure. Let $a^h = (a_1^h, \ldots, a_I^h)$ be the (column) vector of variables describing the K socio-demographic characteristics of the hth household[4] (with $K < H$, where H is the number of households in the population). Finally let $y = c(u^h, p, a^h)$ be a conditional cost function satisfying the usual economic theoretic conditions. Following Lewbel (1985, pp. 3–6), it can be shown that, under an appropriate set of restrictions on the functions $f(.)$ and $g(.)$.

$$c(u^h, p, a^h) = f(y^h, p, a^h) \tag{1}$$

where $y^* = c^*(u^h, p^*)$ and $p^* = g(p, a^h)$, and hence $c^*(u^h, p, a^h)$ can be thought of as a transformation (a 'modification') of the original cost function $c^*(.)$, which, in turn, can be understood as the cost function defined over the intermediate goods (p^* being a vector proportional to the shadow price vector of such goods, whose quantity vector will be denoted by q^*).

In equation (1), the functions $f(.)$ and $g(.)$ (i.e. Lewbel's 'modifying function') permit a full interaction of demographic effects with prices and total expenditure and nest, as special cases, the best-known techniques for introducing demographic variables into demand systems. The present paper focuses on a particular specialization of equation (1), given by the following conditional cost function due to Gorman (1976, pp. 219–221):

$$c(u^h, p, a^h) = c(u^h, p^*) + p'\rho(a^h) \tag{2}$$

and corresponding to equation (1) if $f(.) = y^* - p'\rho(a^h)$ and $g(.) = \Lambda p$ where $\rho(a^h)$ is a vector of demographically varying overheads and Λ is a full matrix of dimension $I \times I$ whose elements are all demographically varying functions. The general procedure embodied in equation (2) will be labelled *extended Gorman* (EG) for reasons that will be apparent below.

In equation (2), production technologies differ among households since intermediate goods prices depend on household characteristics as follows:

$$p_i^* = p_i\left[\lambda_{ii}(a^h) + \sum_{j \neq i} \lambda_{ij}(a^h)(p_j/p_i)\right] \tag{3}$$

In equation (3) input goods prices are *scales* by a good-specific demographic function. The existence of non-zero off-diagonal elements in the matrix emphasizes the possibility of interactions among demographic variables and prices, allowing for jointly demographically varying effects across commodities. In addition, in equation (2), demographically varying fixed costs are added to the demographically scaled cost function. In Gorman's (1976, p. 220) words: 'Why we should be interested in this particular collection of household technologies is another matter. I suppose we might postulate a selfish father who nevertheless found it socially necessary to buy his wife flowers and his children toys, for instance, each time he bought himself a new golfclub.' More prosaically, we could think of small households consuming 'nourishment'

[4] For example, a_1^h could represent family size, a_2^h the age of the head of the household, a_3^h its educational status, and so on.

derived by combining together individual food items, fuel and light, and other input goods in proportions substantially different from those typical of a large household. Alternatively (and more imaginatively) one could think of each intermediate good corresponding to a different household member. To put it in more general terms *extended* procedures are likely to be of considerable practical interest whenever that is a lack of correspondance between marketed goods and consumers' 'wants' *à la* Menger (1871), or, for that matter, consumers' 'objectives' *à la* Hicks (1956). We would like to suggest that this is the rule rather than the exception.

These concepts can be made even clearer by inverting the Marshallian demand corresponding to equation (2), i.e.

$$q = \rho(a^h) + \Lambda'[h(y - p'\rho(a^h), \Lambda p)] \tag{4}$$

where $h(.)$ is the vector of Marshallian demands corresponding to the original ('unmodified') structure of preferences, and solving for $h(.)$ to obtain the household technology implied by equation (2):

$$q^* = h[y - p'\rho(a^h), \Lambda p] = (\Lambda')^{-1}[q - \rho(a^h)] \tag{5}$$

where 'net' input goods are combined together through a linear technology with price independent technical coefficients. Quite clearly, equation (5) suggests that an interesting extension, not attempted in the present paper, could be provided by the case of a rectangular matrix allowing for a number of intermediate goods not equal to the number of inputs. Obviously, in this case we would make use of a generalized inverse of Λ in equation (5).

The cost function (2) nests the best-known procedures for introducing demographic variables into demand systems and a substantial number of specializations can be obtained by imposing successive restrictions on the elements of the Λ matrix, on the elements of the ρ vector, on the interaction between the two.[5]

On the one hand, by restricting ρ to be the null vector we obtain the *extended demographic scaling* specification (Gorman, 1976; (ES))[6] the *demographic scaling* specification (Barten, 1964; (S)), or the single-scale model (Engel, 1985; (E)),[7] depending on whether Λ is assumed to be a full matrix, a diagonal matrix, or an identity matrix times a scalar, respectively. Notice that expressing these restrictions in terms of equation (5) provides further evidence on the inadequacy of non-extended procedures, in that it clearly points out that Barten-like procedures do not allow intermediate goods to be produced by combining market goods. In fact, Barten-type procedures, and, *a fortiori*, Engel's procedure, quite unrealistically imply a diagonal technical coefficient matrix for the household technology. On the other hand, if we allow for the introduction of overheads into the cost function previously defined in terms of demographically scaled prices *à la* Barten, we obtain the Gorman

[5] In turn, the cost function (2) can be generalized by allowing (i) a non-additive interaction between overhead costs and the utility-dependent cost component and (ii) a non-additive interaction between the itemized components of overhead costs. The available evidence (Ray, 1996) suggests the former generalization to be irrelevant.

[6] *Extended demographic scaling*, as defined in the present paper, should not be confused with Lewbel's (1985, pp. 11–13) *generalized scaling* or Ray's (1993) generalized cost scaling. The latter can be understood as a *demographic scaling* with the λ_{ii}'s ($\forall i$) varying with prices and expenditure and not just with demographic variables.

[7] Engel's *single-scale* model reduces almost to nothing the pattern of interaction allowed by equation (1): production technologies are the same across households once deflated by a sort of cost of living index reflecting household needs. Nevertheless Engel's hypothesis of a pattern of demand independent of household characteristics (demographic homotheticity) deserves some attention in the light of the tradition of estimation of *single-scale* models for policy purposes (Deaton and Muellbauer, 1986).

specification (Gorman, 1976; (G), while, setting the matrix equal to the identity matrix, yields the demographic translating specification (Pollak and Wales, 1978; (T)).

3. DEMOGRAPHIC FUNCTIONS AND EFFECTS

As long as particular functional forms are not specified for the demographic functions, the previous section accounts for all general procedures that can be derived by *scaling* and/or *translating* (i.e. introducing overheads into) the original cost function.

However, as Pollak and Wales (1981, p. 1537) point out, as soon as the nature of the demographic functions is specified, a whole new family of so-called reverse procedures can be derived by introducing demographically scaled prices into an already overhead augmented cost function. These procedures (called *reverse Gorman* (RG) and *extended reverse Gorman* (ERG)), would not be distinct in the general case.

The theoretical appeal of such procedures is far from clear in that they imply a cost function in which fixed cost are evaluated at shadow prices, and not at market prices. Nevertheless the previous empirical success of the *reverse Gorman* procedure[8] calls for a closer scrutiny. Consider the *extended reverse Gorman* procedure which it yields the following Marshallian demand:

$$q = \Lambda'\rho(a^h) + \Lambda'[h(y - p'\rho(a^h), \Lambda p)] \tag{6}$$

which differs from equation (4) in that the product $\Lambda'\rho(a^h)$ replaces the vector $\rho(a^h)$.[9] Letting $\lambda_{ij}(a^h) = 0$ $(\forall i \neq j)$ in equation (6) would yield instead the *reverse Gorman* procedure originally proposed by Pollak and Wales (1981).

From equation (6) it is apparent that, for any given demographically varying function $\lambda_{ij}(a^h)$ and $\rho_i(a^h)$, *reverse* procedures show a higher degree of non-linearity than the corresponding *Gorman* procedure, though with the same number of parameters.

In order to clarify this point, consider the parameterization introduced in the empirical estimation in the next section, where the functions $\lambda_{ij}(a^h)$ and $\rho_i(a^h)$ are linear functions of a single demographic variables:

$$\lambda_{ij}(a^h) = \delta_{ij} + \lambda_{ij}a^h \tag{7}$$

$$\rho_i(a^h) = \rho_i a^h \tag{8}$$

where a^h (now a scalar) denotes the number of household components, δ_{ij} is the Kronecker delta and the λ_{ij}'s and the ρ_i's are parameters to be eventually estimated.

The hypothesis of linear functional forms in equations (7) and (8) would appear to be clearly restrictive in that it implies marginal effects independent of family size. This is certainly true for Engel's *single-scale*, as well as for *demographic scaling*, *demographic translating*, *Gorman*, and *extended Gorman*. It is not true, however, for procedures like *reverse Gorman* where the pattern of

[8] Pollak and Wales (1981, p. 1544) find the *reverse Gorman* procedure superior to the *Gorman* in a Generalized CES demand system incorporating a single demographic variable (the number of children in the household) on the basis of the values of their likelihood function. As they admit, though, this is a decision criterion more than a hypothesis test. Bollino and Rossi (1985) reach the same conclusion formally testing, however, *reverse Gorman* against *Gorman*.

[9] To see that *extended Gorman* and *extended reverse Gorman* are not distinct procedures, it suffices to define the translation parameter as in the *extended reverse Gorman* procedure.

interactions (as given by the product $\Lambda'\rho(a^h)$ in equation (6)) is such as to allow quadratic demographic effects even under linear specifications for the $\lambda_{ij}(a^h)$ and $\rho(a^h)$ demographic functions.

This feature appears to be in line with commonsense observations suggesting that fixed costs should bear a non-linear relationship to household size and composition. For instance, once a children's room is set up it usually accommodates more than the first child. In general, it is likely that reverse procedures are more adequate to describe overheads in empirical applications. This feature becomes more evident as the dimension of the Engel space increases, that is, when the model is quadratic in income.

The previous, simple, result provides also an easy way of setting up a formal comparison between *Gorman* and *reverse Gorman*, i.e. two non-nested procedures. With reference to the above linear specification a formal comparison of the two competing hypotheses is possible by specifying a more general hypothesis, incorporating linear *demographic scaling* (equation (7)) and quadratic *demographic translating*, i.e.

$$\rho_i(a^h) = (\rho_i + \eta_i a^h)a^h \tag{9}$$

into the *extended Gorman* procedure. The *extended quadratic Gorman* so derived nests, as special cases, EG ($\eta_i = 0, \forall i$) and ERG ($\eta_i = \Sigma_i \rho_j \lambda_{ij}, \forall i$). Of course, a *quadratic Gorman* (QG) procedure can also be defined correspondingly, nesting *Gorman* ($\eta_i = 0, \forall i$) and *reverse Gorman* ($\eta_i = \rho_i \lambda_{ii}, \forall i$).

Therefore, contrary to Pollak and Wales (1981), the introduction of EQG and QG allows an unambiguous ranking of procedures. In this respect, the present paper hopes to shed some light on the issue of non-linearity of demographic effects by considering three separate cases: linear demographic functions and linear demographic effects (e.g. *Gorman* and *extended Gorman*), linear demographic function and non-linear demographic effects (e.g. *reverse Gorman* and *extended reverse Gorman*), and non-linear demographic functions and effects (e.g. *quadratic Gorman*, *extended quadratic Gorman*).

4. ESTIMATION AND EMPIRICAL RESULTS

This section estimates and compares the general procedures described in the previous section, using each of them to incorporate a single demographic variable, family size, into a Generalized Quadratic Almost Ideal demand system[10] (Bollino, 1987; Banks, Blundell, and Lewbel, 1997) of rank 3 (Lewbel, 1991). The Generalized QAI demand equations (in share form and incorporating our most general procedure, i.e. EQG) are given by:

$$w_i = \sum_j \phi_j(a^h)\lambda_{ij}(a^h)p_i/y + (\bar{y}/y)\sum_j \lambda_{ij}(a^h)\left\{\alpha_j + \sum_k \gamma_{jk}\ln p_k^* + [\beta_j + \psi\ln(\bar{y}/P)]\ln(\bar{y}/P)\right\} \tag{10}$$

where

$$w_i = p_i q_i / y_i,\ \bar{y} = y - \sum_k \phi_k(a^h)p_k,\ \phi_j(a^h) = \phi_j + \delta_{ij}[\rho_j(a^h)/\lambda_{ij}(a^h)],\ p_i^* = \sum_j \lambda_{ij}(a^h)p_j$$

[10] Which may be regarded as the least intrusive quadratic modification of a demand system that maintains integrability.

Table I. Estimation results (no. of observations: 1729)

System	No. of parameters	Log-likelihood
Linear Expenditure Systems (LES)	5	−2215·96
Generalized Almost Ideal (GAI)	10	−1368·47
Generalized Quadratic Almost Ideal (GQAI)	11	−1353·52
Single-scale (E)	12	−1270·01
Demographic scaling (S)	14	−633·04
Demographic translating (T)	14	−914·29
Quadratic translating (QT)	17	−625·18
Gorman (G)	17	−542·95
Reverse Gorman (RG)	17	−566·24
Quadratic Gorman (QG)	20	−492·56
Reverse quadratic Gorman (RQG)	20	−490·31
Extending scaling (ES)	20	−488·45
Extended Gorman (EG)	23	−409·64
Extended reverse Gorman (ERG)	23	−460·87
Extended quadratic Gorman (EQG)	26	−370·41
Unrestricted (U)	66	67·97

the functions $\lambda_{ij}(a^h)$ and $\rho_j(a^h)$ are given by equations (7) and (9), respectively, P is the usual Almost Ideal price aggregator to be evaluated according to *Stone's* approximation ($\Sigma_k w_k \ln p_k$) and the following restrictions hold:[11]

$$\sum_k \alpha_k = 1; \quad \sum_j \gamma_{kj} = \sum_k \gamma_{jk} = \sum_k \beta_k = 0; \quad \gamma_{kj} = \gamma_{jk}$$

Empirical estimates are based on a data set extracted from the ISTAT annual household survey, for the period 1973–1992. Budget expenditures are grouped by expenditure classes and family size for a total of 1815 observations; exclusion of non-reporting and/or meaningless cells reduces the sample size to 1729 (see Appendix). It is relevant to emphasize that household size is the only demographic variable available from published data. This informational limitation does not allow us to fully account for systematic heterogeneity in preferences. Furthermore, it must be noted that model adequacy can be significantly affected by the high level of data aggregation in terms of aggregation both across households due to the fact that the published data are in cell means and across goods. The price vector has been derived from the corresponding consumption breakdown available from NIA.[12] In order to limit the computational complexity, only three categories have been analyzed: 'food', 'housing and fuels', and 'miscellaneous'.[13]

[11] Abstracting from demographic effects, if $\varphi_i = \psi = 0$, equation (10) reduces to the original Deaton–Muellbauer AI. It is important to note the role of committed quantities in the QAI system. In this respect the Generalized QAI is different from budget share translating as proposed by Lewbel (1985) and there is no need for parametric restrictions on the translating parameter in order to satisfy the usual demand theory properties (adding-up and homogeneity).
[12] The use of household specific weights provides a (useful by limited) price variability across households.
[13] The 'miscellaneous' item includes clothing, health care, transportation, communications and all other goods and services.

The stochastic specification of each demand system is assumed to be additive in the share form of each equation, where the covariance matrix of the disturbances is singular in observance of the adding up constraint. All estimates discussed below are obtained with a non-linear FIML procedure with observations weighted by the relative frequency of households in each expenditure class. The estimated covariance matrix in heteroscedasticity consistent.

Table I presents the log likelihood of the estimated systems and the number of estimated parameters. On the basis of Table I likelihood ratio values (denoted as $\chi^2_{(n)}$, with n indicating degrees of freedom) for all admissible direct nested tests are immediately derived.

Apart from allowing an assessment of the systematic influence of demographic characteristics on consumption patterns, three basic questions concerning consumer behaviour can be answered with the help of Table I. First, which representation of consumer preferences appears to be preferred in a statistical sense? Second, is a non-linear specification for demographic effects significantly superior to a corresponding linear specification? If so, does the answer depend on the non-linearity of the underlying demographic function or on the interaction of the demographic effects allowed for by the different procedures?

The upper part of the table, in agreement with other studies on other data sets, shows that the Generalized Quadratic Almost Ideal model without demographic characteristics is statistically superior to rank 2 models like the Linear Expenditure System (LES) and the Generalized Almost Ideal (GAI) model ($\chi^2_{(6)}$=1724·88 and $\chi^2_{(1)}$=29·9 for LES and GAI, respectively).

In agreement with all the previous work on the subject, it can be immediately concluded that there exists a significant effect of demographic characteristics. By inspection of the upper part of Table I it is clear that the (generalized quadratic) AI without demographics is soundly rejected even when compared with the simplest demographically augmented demand system (i.e. Engel's *single-scale*, $\chi^2_{(1)}$=167·02). In turn, the *single-scale* model is rejected against the Barten's *demographic scaling* ($\chi^2_{(2)}$=1273·94), thus giving support to the hypothesis of commodity specific demographic effects.

As far as *demographic scaling*, and *demographic translating* are concerned, the two hypotheses are non-nested. A formal test is possible, though, by comparing S and T to more general procedures, like G and RG, incorporating them as special cases. As it turns out, contrary to Pollak and Wales (1981, pp. 1543–1545) and in accordance with Ray's (1985) findings, *demographic scaling* is rejected against the more general hypotheses ($\chi^2_{(3)}$=180·18 and $\chi^2_{(3)}$=133·60, for G and RG, respectively) along with demographic translating ($\chi^2_{(3)}$=742·68 and $\chi^2_{(3)}$=696·10, for G and RG, respectively). It then becomes of paramount importance to compare formally the two non-nested hypotheses given by G and RG since such comparison may provide information about the non-linearity of demographic effects along with information about the relative showing of different demographic procedures.

As suggested in Section 3, this comparison can legitimately be formed in terms of a more general hypothesis given by QG. In this respect, both G and RG do not compare favourably with QG, though this is less so for G ($\chi^2_{(3)}$=100·78 and $\chi^2_{(3)}$=147·36, for R and RG, respectively). In addition, although not directly comparable with Pollak and Wales' (1981) estimation of a quadratic *demographic scaling*, the mixture of linear *demographic scaling* and quadratic *demographic translating* given by QG suggests that existence of non-linear demographic effects based on non-linear demographic functions.

A partially novel result in empirical demand analysis is shown in the lower part of Table I, where we provide a systematic estimation and comparison for all four different *extended* procedures, i.e. ES, EG, ERG and EQG. The first thing to note is that *extended* procedures

Table II. Parameter estimates, estimated budget shares and estimated elasticities (year: 1992, system: *extended quadratic Gorman*, No. of observations: 1729)

Estimation results	$i = 1$	$i = 2$	$i = 3$
Parameter Estimations			
α_i	0·834 (0·020)	0·164 (0·009)	0·002 (0·022)
γ_{i1}	0·484 (0·010)	−0·212 (0·008)	−0·272 (0·011)
γ_{i2}	−0·212 (0·008)	0·049 (0·008)	−0·163 (0·010)
γ_{i3}	−0·272 (0·011)	−0·163 (0·010)	0·109 (0·016)
β_i	−0·591 (0·015)	0·122 (0·007)	0·468 (0·015)
ψ	−0·194 (0·004)	−0·194 (0·004)	−0·194 (0·004)
φ_i	−0·740 (0·033)	−0·028 (0·014)	0·112 (0·018)
ρ_i	0·077 (0·009)	0·042 (0·005)	−0·036 (0·007)
η_i	−0·023 (0·002)	−0·006 (0·001)	−0·001 (0·001)
λ_{i1}	0·084 (0·005)	−0·011 (0·005)	−0·102 (0·010)
λ_{i2}	−0·064 (0·010)	−0·009 (0·013)	0·077 (0·013)
λ_{i3}	−0·000 (0·000)	−0·000 (0·000)	−0·128 (0·001)
R^2	0·985	0·943	–
Estimated mean budget shares			
w_i	0·391	0·268	0·341
Estimated mean elasticities			
Family size I			
e_i	0·313	1·115	1·811
e_{i1}	−0·162	−0·978	0·827
e_{i2}	−0·473	−0·844	0·202
e_{i3}	0·635	1·822	−1·030
Family size 2			
e_i	0·432	1·030	1·689
e_{i1}	−0·451	−1·005	1·024
e_{i2}	−0·455	−0·714	0·138
e_{i3}	0·906	1·719	−1·162
Family size 4			
e_i	0·664	0·893	1·501
e_{i1}	−1·069	−1·190	1·595
e_{i2}	−0·466	−0·575	0·148
e_{i3}	1·535	1·764	−1·743

Note: $i = 1$ indicates 'food', $i = 2$ indicates 'housing and fuel', $i = 3$ denotes 'miscellaneous goods and services'.

significantly generalize their specialized counterparts. The null defined by S, G, RG, and QG are all rejected against the alternatives given by ES ($\chi^2_{(6)}$=289·19), EG ($\chi^2_{(6)}$=266·62, ERG ($\chi^2_{(6)}$=210·74), and EQG ($\chi^2_{(6)}$=244·30), respectively. Therefore, it seems safe to conclude that the available evidence strongly suggests that prices interact with demographic variables in a much richer way than usually allowed for in the previous literature. Again, *demographic scaling* makes a poor showing against *reverse Gorman* even in the new *extended* setting ($\chi^2_{(3)}$=55·16). Finally, the previous result concerning the non-linearity of demographic effect appears to receive further support in the extended case, since the *extended reverse Gorman* (ERG) turns out to be soundly rejected when compared with an *extended quadratic Gorman* (EQG, $\chi^2_{(3)}$=180·92).

It is relevant to note that the statistical comparisons presented in Table I are based on likelihood ratio tests whose reliability requires normally distributed errors, we investigated the behaviour of the heteroscedasticity consistent residuals plotting the non-parametric kernel

Table III. Estimated linear household technologies (year: 1992, system: *extended quadratic Gorman*)

Linear household technologies	$i = 1$	$i = 2$	$i = 3$
Family size 1			
λ^{i1}	0·923	0·060	0·000
λ^{i2}	0·010	1·010	0·000
λ^{i3}	0·107	−0·082	1·147
Family size 2			
λ^{i1}	0·858	0·112	0·000
λ^{i2}	0·019	1·021	0·000
λ^{i3}	0·231	−0·181	1·344
Family size 4			
λ^{i1}	0·755	0·201	0·000
λ^{i2}	0·034	1·046	0·000
λ^{i3}	0·610	−0·493	2·049

Note: $i = 1$ indicates 'food', $i = 2$ indicates 'housing and fuel', $i = 3$ denotes 'miscellaneous goods and services'.

densities of the disturbances and computing both the *Bera* and *Jarque* and the Skewness/Kurtosis tests. While non-normality is not apparent from the graphical analysis, the tests' results show that non-normality is due mainly to a high kurtosis, as one may expect considering that the residuals have not been weighted for the relative frequency of households in each expenditure class. Correcting for this error structure restores normality. Nonetheless, we complemented the results of the statistical comparisons of the Extended Quadratic Gorman model versus the nested Extended Gorman and Reverse Gorman specification based on likelihood ratio tests also performing the *Wald* test. The test rejects the null hypothesis that the EQG parameters are jointly zero both if compared to the EG and ERG demographic specification ($\chi^2_{(3)}$=234·99 and $\chi^2_{(3)}$=205·43, for EG and ERG, respectively). These results are consistent with the results obtained using the likelihood ratio test.

In the light of the above results, estimated elasticities for the extended systems have been computed at all family expenditure levels for different family sizes. For all systems it should be noted that in general the estimated parameters correspond to 'well-behaved' preferences in all price–expenditure demographics situations, satisfying concavity and Slutzky negativity conditions. This regularity is additional evidence of model adequacy of extended specifications.[14]

In Table II, we report for EQG compensated price elasticities, expenditure elasticities, budget shares computed at average expenditure level, together with a measure of goodness-of-fit of budget shares equations and estimated parameter values.

A glance at Table II shows that, as expected, 'food' is a necessity at all family sizes, 'housing and fuel' becomes one only when (presumably) children are present and 'miscellaneous' shows as a luxury, although less so as family size increases. 'Food' exhibits the lowest gross own price elasticity in absolute value for family sizes smaller or equal to two, while 'housing and fuel' shows as the most inelastic item for family sizes greater than two. Finally, 'miscellaneous' presents the highest own price elasticity. Moreover, computing the figures of Table II at different levels of expenditure (not reported here), suggest that for given family size, the shares of goods other than good increase with expenditure. In addition, own price elasticities decrease, in absolute value,

[14] Detailed information about parameter estimates and elasticities for all demand systems are available from the authors on request. Violations of regularity conditions are generally confined to extreme socio-demographic combinations.

with total expenditure. These results are well in line with previous estimates of Italian consumption patterns (Bollino and Rossi, 1985).

Finally, Table III provides evidence of the implied demographically varying household technology and addresses the issue of economies of scale in consumption. In particular, it reports the estimated matrix $(\Lambda')^{-1}$ (with typical element denoted by λ^{ij}) whose rows transform input goods (not of overhead) into intermediate goods, which, for the sake of argument, could be labelled as 'nourishment', 'shelter' and 'social relations', respectively. A few things are worth noting. As family size increases, producing 'nourishment' requires increasingly less 'food' and more 'housing and fuel', while producing 'social relations' requires increasingly larger amounts of 'food' and 'miscellaneous goods and services' and increasingly less 'housing and fuel'. Moreover, summing the λ^{ij}'s over j provides additional evidence of the issue of economies of scale in that $\Sigma_j \lambda^{ij}$ corresponds to the amount of intermediate goods derived from one unit of each input good (net of overhead). As it turns out, as family size increases,[15] the amount of 'nourishment' produced by one unit of 'food', 'housing and fuel' and 'miscellaneous' remains practically the same (and very near one). Mild economies of scale show in the production of 'shelter' for which $\Sigma_j \lambda^{ij}$ goes from 1·020 (for family size one) to 1·081 (for family size four), while the production of 'social relations' appears to be dominated by economies of scale since $\Sigma_j \lambda^{ij}$ ranges from 1·172 (for family size one) to 2·166 (for family size four).

However, general as they are, linear household technologies (with or without non-linear demographic function) are still far from providing a satisfactory representation of consumer behaviour. As the last row of Table I shows, when compared with an unrestricted (unpooled) generalized QAI demand system, even the *extended quadratic Gorman* fails to stand up, its implied restrictions being overwhelmingly rejected. This negative result suggests that, notwithstanding the clear improvement provided by non-linear structural approaches to incorporate demographics into demand systems, these are still a long way from explaining actual behaviour.

5. CONCLUSIONS

In the present paper, the best-known general procedures for incorporating demographic variables into complete demand systems have been discussed, estimated and compared. A class of general procedures belonging to Gorman's family of 'general linear household technologies' has been introduced, discussed, and estimated. These *extended* procedures allow for varying degrees of non-linearity of demographic functions and demographic effect, and imply that marketed goods (net of overheads) are combined together by households through a linear price independent technology in order to produce intermediate goods and generate utility.

Comparisons have been based on a Generalized Quadratic Almost Ideal demand system incorporating a single demographic variable (family size) estimated and tested on Italian household budget data for the years 1973–1992.

Use of normal statistical tests in both nested and non-nested cases has allowed a formal ranking of the procedures. Engel's adult *equivalent model* made the weakest showing while *extended* procedures (like the *extended quadratic Gorman*) dominated the corresponding specialization. Furthermore data showed clear evidence of non-linear demographic effects,

[15] The notion of a change in family size is, to say the least, arbitrary since an additional individual is not necessarily similar to the previous one (Muellbauer, 1977, p. 470).

originating from demographic functions. However, whatever their degree of generality, household technologies have never proved capable of accounting fully for the heterogeneity of consumption data. We deem that the addition of other demographic characteristics, using individual household level data, or the explicit modelling of the information available to the household in the decision-making process (and usually not available to the researcher) might further improve the explanatory power of the structure.

APPENDIX

Expenditure surveys were undertaken by the Italian Central Statistical Office (ISTAT) in 1968 and carried out regularly in subsequent years. However, in 1973 the methodology of the survey went through a series of major revisions. Therefore, results refer only to the homogeneous series of surveys starting in 1973 and used in this paper. The survey covers all private consumers' expenditures including 'self-consumption' of agricultural products, imputed rents and consumer goods received as wages. However, it does not cover purchases of houses or land as well as expenditure related to production.

As far as consumer durables are concerned, expenditures are recorded at the time of purchases. Hence, for example, payments by instalments are not accounted for. The survey's central unit is the household defined as a set of persons living together characterized by the common use of their incomes. The household, as defined in the surveys, includes servants and nurses.

The methodology of the survey can be briefly described as follows. In a first stage 700 towns and cities are selected and divided into two groups: group 1 includes capitals of provinces and towns with at least 50,000 inhabitants, group 2 includes the rest. A predetermined number of households in towns in group 1 are surveyed each month of the year while towns in group 2 are divided into three further subgroups and each subgroup participates the survey in the first, second or third month of each quarter of the year. Summarizing, each month approximately 320 towns are surveyed: 140 of them belonging to group 1 above and 180 to group 2. In a second stage 3000 households per month are randomly selected from the registrar's offices of the selected towns, thus giving a total of 36,000 households surveyed in the year. The rate of participation in the survey has been approximately estimated around 85%. The initial results are then related to the population by ISTAT.

ACKNOWLEDGEMENTS

This is a fully revised version of a previous paper by C.A. Bollino and N. Rossi, circulated under the title 'The interaction of family size with prices and expenditure in demand systems'. The authors wish to thank A.P. Barten, R. Blundell, A. Lewbel and two anonymous referees for their comments and suggestions. Remaining errors are the authors' own.

REFERENCES

Banks J, Blundell R, Lewbel A. 1997. Quadratic Engel curves and consumer demand. *Review of Economics and Statistics* **79**: 527–539.

Barten AP. 1964. Family composition, prices and expenditure patterns. In *Econometric Analysis for National Planning*, Hart PE, Mills G, Whittaker JK (eds). Butterworth: London.

Bollino CA.1987. GAIDS, A generalized version of the Almost Ideal Demand System. *Economics Letters* No. 23

Bollino CA, Rossi N. 1985. Teoria e problemi di stima di effetti demografici nell'analisi della domanda: applicazioni al caso italiano. *Contributi all'Analisi Economica* **1**: 69–99.

Browning M. 1992. Children and household economic behaviour. *Journal of Economic Literature* **30**: 1434–1475.

Deaton AS, Muellbauer J. 1986. On measuring child cost: with applications to poor countries. *Journal of Political Economy* **94**: 720–744.

Engel E. 1985. Die Lebenskosten Belgischer Arbeiter-Familien Fruher Und Jetzt. *International Statistical Institute Bulletin* **9**: 1–74.

Gorman WM. 1976. Tricks with utility functions. In *Essays in Economic Analysis*, Artis MJ, Nobay AR (eds). Cambridge University Press: Cambridge.

Hicks JR. 1956. *A Revision of Demand Theory*, . Oxford University Press: Oxford.

Lewbel A. 1985. A unified approach to incorporating demographic or other effects into demand systems. *Review of Economic Studies* **52**: 1–18.

Lewbel A. 1986. Additive separability and equivalent scales. *Econometrica* **54**: 219–222.

Lewbel A. 1991. The rank of demand systems: theory and non-parametric estimation. *Econometrica* **59**: 711–730.

Menger C. 1871. *Grundsatze der Volkwirtschaftslehre*, (Italian translation: *Principi di Economia Politica*. UTET: Torino 1976).

Muellbauer J. 1977. Testing the Barten model of household composition effects and the cost of children. *Econometrica* **54**: 657–678.

Muellbauer J. 1980. The estimation of the Prais–Houthakker model of equivalence scales. *Econometrica* **48**: 153–176.

Nelson J. 1992. Methods of estimating household equivalence scales: an empirical investigation. *Review of Income and Wealth* **38**: 295–310.

Pollak RA, Wales TJ. 1978. Estimation of complete demand systems from household budget data: the linear and quadratic expenditure systems. *American Economic Review* **68**: 348–359.

Pollak RA, Wales TJ. 1980. Comparison of the quadratic expenditure system and the translog demand system and alternative specifications of demographic effects. *Econometrica* **48**: 595–612.

Pollak RA, Wales TJ. 1981. Demographic variables in demand analysis. *Econometrica* **49**: 1533–1551.

Prais S, Houthakker H. 1955. *The Analysis of Family Budgets*, . Harvard University Press: Cambridge, MA.

Ray R. 1985. A nested test of the Barten model of equivalence scales. *Economics Letters* **17**: 411–412.

Ray R.1993. Nesting demographically extended flexible demand functional forms. UBC DP No. 93–21

Ray R. 1996. Demographic variables in demand systems: the case for generality. *Empirical Economics* **21**(2): 307–315.

Rothbarth E. 1943. Note on a method of determining equivalent income for families of different compositions. In *War-time Pattern of Saving and Spending*, Madge C (eds). Cambridge University Press: Cambridge.

Tsakloglou P. 1991. Estimation and comparison of two simple models of equivalence scales for the cost of children. *Economic Journal* **101**: 343–357.

