Quadratic Engel Curves and Household Equivalence Scales: the Case of Italy 1985 - 1994

Gianni Betti
Department of Statistics
London School of Economics
Houghton Street
London WC2A 2AE, UK

September 15, 2000

Abstract

This paper presents equivalence scales based on a complete demand system that is consistent with the observed expenditure patterns. Preliminary non-parametric analysis, based on kernel estimators and rank-three tests, shows clear evidence of nonlinearities in the Engel curves. A quadratic specification of the underlying Engel curves is utilised: this is the Quadratic Almost Ideal Demand System. The pooled data from the Italian Household Budget Survey 1985 - 1994 is utilised to compare models and to calculate equivalence scales for the preferred model.

Keywords: quadratic Engel curves, equivalence scales, rank three demand systems.

1 Introduction

In poverty analysis, income distribution (or total consumption[(1)]), plays a fundamental role. Whatever variable is chosen as welfare measurement, there is the necessity to derive a new variable to perform a comparison among households; in fact, families vary in size and structure, and the per-capita income (or consumption) is not an adequate measure, because it ignores the economies of scale present in any household.

Equivalence scales are economic index numbers, which discount household income according to some household characteristics. Although equivalence scales are considered by one consent a necessary tool, there is no unanimity in the way they have to be calculated. Buhman *et al* (1988) and Hagenaars *et al* (1994) present a broad classification: i) normative or social security scales; ii) subjective scales; iii) scales based on demand models. Discussion on which approach must be followed, goes beyond the scope of this paper; the method proposed here belongs to the third and most common group. In particular this paper aims at calculating equivalence scales based on complete demand systems. In view of that, equivalence scales are defined following Deaton and Muellbauer (1980), as the relative amount two different types of households spend in order to reach the same level of economic utility:

$$e(u, \mathbf{p}, \mathbf{z}, \mathbf{z}^r) = \frac{C(u, \mathbf{p}, \mathbf{z})}{C(u, \mathbf{p}, \mathbf{z}^r)} \qquad (1)$$

where $C(.)$ is a classical cost function, u is the utility level, $\mathbf{p}$ is the vector of prices, $\mathbf{z}$ and $\mathbf{z}^r$ are the vectors of demographic characteristics of, respectively, a given household and the reference household.

In this paper we also aim at testing the nonlinearities in the Engel curves[(2)] for seven groups of commodities from a pooled data set based on the Italian Household Budget Survey, conducted by the Italian Office of Statistics (ISTAT) from 1985 to 1994, and called *Indagine sui Consumi delle Famiglie* (BF).

The paper is composed of five sections. Section two develops the complete demand systems here utilised. The starting point is the well known model presented by Deaton and Muellbauer (1980). This simple model is extended introducing nonlinearities in the Engel curves, transformation of the concavity of the price kernel, and socio-economic variables.

[(1)] The choice between consumption and income is crucial, but beyond the scope of this paper.
[(2)] An Engel curve relates expenditure on a given commodity to total expenditure by a household.

Section three fully describes the data set from the Italian Household Budget Surveys 1985 - 1994. A preliminary analysis on Engel curves conducted utilising kernel nonparametric regression and rank-three tests shows a clear presence of nonlinearities in the curves. Section four presents the empirical analysis. Following the same order as the theoretical section two, five models are estimated and compared. One model is chosen and equivalence scales are calculated for different households varying in size and composition. The concluding remarks in section five end the paper.

2 Demand analysis and demographic variables

In the economic theory of consumer demand the basic behavioural hypothesis is that households choose a basket of goods that is preferred to all the other baskets that they can afford. Let $u = v(\mathbf{q}, \mathbf{p})$ the households utility function, $\mathbf{q}$ the vector of consumer goods and $\mathbf{p}$ the vector of corresponding prices. The consumer choices follow the maximization of the utility function, subject to the budget constraint $\sum_{i=1}^{n} q_i p_i = \mu$; the quantity $q_i(\mu, \mathbf{p})$is known as Marshallian demand.

The dual problem consists in the minimization of the total expenditure subject to a fixed value of utility u; the yields are the Hicksian or compensated demand functions $h_i(u, \mathbf{p})$. Since the two approaches are equivalent, the following equation holds:

$$\mu = \sum q_i(\mu, \mathbf{p}) p_i = \sum h_i(u, \mathbf{p}) p_i = C(u, \mathbf{p}) \tag{2}$$

Muellbauer (1976) defines as Price-Independent Generalised Logarithmic (PIGLOG) those demand systems having expenditure shares that are linear in logarithm of total expenditure. The Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980) belongs to PIGLOG category.

The indirect utility function of the initial model AIDS is defined as follows:

$$v(x, \mathbf{p}) = \frac{\ln x}{b(\mathbf{p})} \tag{3}$$

where $x = \frac{\mu}{a(\mathbf{p})}$is the expenditure in real terms;

$\ln a(\mathbf{p}) = \alpha_0 + \sum \alpha_k \ln p_k + \frac{1}{2} \sum \sum \gamma_{kj} \ln p_k \ln p_j$ and $\ln b(\mathbf{p}) = \sum \beta_k \ln p_k$ are price indices.

The associated cost function results to be:

$$\ln C(u, \mathbf{p}) = \alpha_0 + \sum \alpha_k \ln p_k + \frac{1}{2} \sum \sum \gamma_{kj} \ln p_k \ln p_j + \prod p_k^{\beta^k} u \tag{4}$$

and from Roy's identity one can obtain the budget shares as a function of total expenditure, μ, and prices $\mathbf{p}$:

$$w_i = \frac{\partial \ln a(\mathbf{p})}{\partial \ln p_i} + \frac{\partial \ln b(\mathbf{p})}{\partial \ln p_i} \ln x = \alpha_i + \sum \gamma_{ij} \ln p_j + \beta_i \left(\ln \mu - \ln a(\mathbf{p})\right). \tag{5}$$

This simple Engel curve can be extended in three different directions, introducing:

- nonlinearities in the Engel curves;
- transformation of the price kernel concavity;
- socio-demographic variables.

2.1 Rank three demand systems

Lewbel (1991) defines the rank of a demand system as the dimension of the space spanned by the implicit Engel curves; by this definition the rank of AIDS (5) is equal to two.

Banks, Blundell and Lewbel (1997) starts from PIGLOG extending to a simple general form of demand systems consistent with nonlinear Engel curves:

$$w_i = A_i(\mathbf{p}) + B_i(\mathbf{p}) \ln x + C_i(\mathbf{p}) g(x) \tag{6}$$

where $A_i(\mathbf{p})$, $B_i(\mathbf{p})$, $C_i(\mathbf{p})$ and $g(x)$ are differentiable functions.

The rank of demand system (6) corresponds to the rank of the N x 3 matrix of Engel curve coefficients $[A(\mathbf{p})\ \ B(\mathbf{p})\ \ C(\mathbf{p})]$, whose maximum rank is three.

The uniqueness of a rank three demand system form (6) is guaranteed by a theorem in Banks, Blundell and Lewbel (1997, p.532):

Theorem 1 *All exactly aggregable demand systems in the form of equation (6) that are derived from utility maximization either have:*

$$C_i = d(\mathbf{p}) B_i(\mathbf{p}) \tag{7}$$

for some function (so the rank is less than 3), or they are rank 3 quadratic logarithmic budget share systems having indirect utility function of the form:

$$\ln v(\mu, \mathbf{p}) = \left\{ \left(\frac{\ln \mu - \ln a(\mathbf{p})}{b(\mathbf{p})} \right)^{-1} + \phi(\mathbf{p}) \right\}^{-1} \tag{8}$$

where the term $\{\ln \mu - \ln a(\mathbf{p})\} / b(\mathbf{p})$ is the indirect utility function of a PIGLOG demand system (i.e, a system with budget shares linear in log total expenditure), and the extra term $\phi(\mathbf{p})$ is a differentiable, homogenous function of degree zero of prices $\mathbf{p}$.

Banks, Blundell and Lewbel (1997) chooses a linear and homogenous specification of $\phi(\mathbf{p})$: $\ln \phi(\mathbf{p}) = \sum \phi_k \ln p_k$, and calls the demand system Quadratic Almost Ideal (QUAIDS).

From this model the corresponding budget shares are:

$$w_i = \alpha_i + \sum \gamma_{ij} \ln p_j + \beta_i \ln x + \frac{\phi_i}{b(\mathbf{p})} (\ln x)^2. \tag{9}$$

2.2 CES transformation

A transformation of the price kernel concavity of type CES (Constant Elasticity of Substitution) modifies the functional form $a(\mathbf{p})$ as follow:

$$\ln a(\mathbf{p}) = \alpha_0 + \frac{1}{(1-\sigma)} \ln \left(\sum \alpha_k \ln p_k^{(1-\sigma)} \right) + \frac{1}{2} \sum \sum \gamma_{kj} \ln p_k \ln p_j \tag{10}$$

which corresponds to the Generalised Almost Ideal Demand System (GAIDS) by Lancaster and Ray (1998); in order to be coherent with the terminology we call it CES_QUAIDS. By Roy's identity the budget shares are:

$$w_i = \frac{\alpha_i p_i^{(1-\sigma)}}{\sum \alpha_k \ln p_k^{(1-\sigma)}} + \sum \gamma_{ij} \ln p_j + \beta_i \ln x + \frac{\phi_i}{b(\mathbf{p})} (\ln x)^2 . \tag{11}$$

2.3 Demographic variables

The demand systems introduced so far, make the assumption that households behave in the same manner in choosing the basket of goods, in order to maximise their economic utility. But households differ in size and composition; for example, it is easy to imagine that preferences of a young couple differ a lot from those of an elderly one.

For this reason over the last four decades economic modelling has aimed at introducing demographic variables into the utility and, indirectly, into the cost function. Barten (1964) considers the utility function associated with the household demographic characteristic to be:

$$u = v \left(\frac{q_1}{m_1(\mathbf{z})}, \frac{q_2}{m_2(\mathbf{z})}, ..., \frac{q_n}{m_n(\mathbf{z})} \right) \tag{12}$$

which corresponds to the cost function:

$$C^1(u, \mathbf{p}, \mathbf{z}) = C\left[u, p_1 m_1(\mathbf{z}), p_2 m_2(\mathbf{z}), ..., p_n m_n(\mathbf{z})\right] \tag{13}$$

where $m_i(\mathbf{z})$ is the equivalence scale for the particular good i; all the $m_i(\mathbf{z})$ are equal to unity in the case of the reference household. This model is known in literature as *Demographic Scaling*. Although the model is more general than those without demographic variables, there is a problem in the evaluation of equivalence scales in goods which are not consumed in the reference household (for example child food when the reference household is the couple).

Gorman (1976) modifies the previous model by introducing a new term directly into

the cost function:

$$C^2(u,\mathbf{p},\mathbf{z}) = C\left[u, p_1 m_1(\mathbf{z}), p_2 m_2(\mathbf{z}), ..., p_n m_n(\mathbf{z})\right] + \sum p_k d_k(\mathbf{z}) \tag{14}$$

where the added last term on the right side represents the fixed cost associated with the demographic characteristic vector $\mathbf{z}$.

Pollak and Wales (1978) proposes a simpler method, termed *Demographic Translating*, which corresponds to the cost function:

$$C^3(u,\mathbf{p},\mathbf{z}) = C(u,\mathbf{p}) + \sum p_k d_k(\mathbf{z}). \tag{15}$$

The authors state that the Gorman (14) is the general model and it includes the *Demographic Translating* specification (when the $m_i(\mathbf{z})$ are unity) and the *Demographic Scaling* specification (when the $d_k(\mathbf{z})$ are all zero). The Gorman model consists of the following operations of scaling and translating of the original demand system (the one without demographic characteristics).

Pollak and Wales (1981) inverts the order of the above operations, obtaining a new model called *Reverse Gorman*, whose corresponding cost function is:

$$C^4(u,\mathbf{p},\mathbf{z}) = C\left[u, p_1 m_1(\mathbf{z}), p_2 m_2(\mathbf{z}), ..., p_n m_n(\mathbf{z})\right] + \sum p_k m_k(\mathbf{z}) d_k(\mathbf{z}) \tag{16}$$

A proposal, which is alternative to the previous ones, is due to Ray (1983). In this model, called *Price Scaling*, the term included the demographic variables is multiplicative with respect to the original cost function:

$$C^5(u,\mathbf{p},\mathbf{z}) = C(u,\mathbf{p}) m(\mathbf{p},\mathbf{z}). \tag{17}$$

Lewbel (1985) presents a unifying approach incorporating demographics or other effects into demand systems; such an approach is based on the technique of cost function modification, using a general transformation $C^6\left(u,\mathbf{p},\mathbf{z}\right) = f\left[C(u, h\left(\mathbf{p},\mathbf{z}\right)), \mathbf{p}, \mathbf{z}\right]$. Lewbel gives a set of restrictions for proving that Barten, Gorman and both Pollak and Wales models are special cases.

Bollino and Rossi (1989) presents an extension to the *Reverse Gorman*, where the relative prices are strictly dependent on the scaling coefficients:

$$C^7(u,\mathbf{p},\mathbf{z}) = C\left(u,\mathbf{p}^*\right) + \sum p_k^* d_k(\mathbf{z}) \tag{18}$$

where

$$p_k^* = p_k \left[m_k(\mathbf{z}) + \sum_{j \neq k} m_j(\mathbf{z}) \left(\frac{p_j}{p_k} \right) \right]. \tag{19}$$

An equivalence scale is defined exact (ESE, Equivalence Scale Exactness, Blackorby and Donaldson, 1989), when it is independent of the utility level. For this reason that property is also known in literature as IB (Independent from the Base, Lewbel, 1989a). Analytically the ESE or IB property can be formulated as follow:

$$ES = \frac{C(u, \mathbf{p}, \mathbf{z})}{C(u, \mathbf{p}, \mathbf{z}^r)} = \frac{G(u, \mathbf{p}) m(\mathbf{p}, \mathbf{z})}{G(u, \mathbf{p}) m(\mathbf{p}, \mathbf{z}^r)} = \frac{m(\mathbf{p}, \mathbf{z})}{m(\mathbf{p}, \mathbf{z}^r)}. \tag{20}$$

Among the seven models $C^1 - C^7$ introduced so far, only the *Price Scaling* satisfies the ESE property, without imposing any a priori restrictions on the demand system specification.

3 The data set

This paper aims at verifying which complete demand system is appropriate for Italy; an in depth analysis is conducted to test the linearity of Engel curves using both non-parametric regression and demand rank tests. This preliminary analysis confirms that quadratic demand systems are appropriate.

3.1 The Italian ISTAT Household Budget Survey 1985 - 1994

This study is based on the Italian Household Budget Survey conducted by the Italian Office of Statistics (ISTAT), from 1985 to 1994, and called *Indagine sui Bilanci delle Famiglie* (BF). Each cross-section of survey is composed of more than 32,000 records corresponding to households interviewed. All samples are independent from each other and they are drawn by a two-stage sampling. Primary sampling units, PSUs, are the municipalities (*comuni*), and the ultimate sampling units USUs are the households.

Municipalities are divided into: a) self-representative (AR), referring chief towns of province (96) or other municipalities with more than 50,000 inhabitants (52); b) non self-representative (NAR), referring municipalities with less than 50,000 inhabitants. PSUs are stratified in the following manner: every AR municipality forms a stratum and it is included in the survey every month. NAR municipalities are stratified for: region, altimetric zone and main economic activity; from each stratum, three municipalities are

drawn, each of them is in the survey for one in three months.

The sample is so composed of all AR municipalities (148) and 134 NAR municipalities.

Once the PSUs have been drawn, the USUs or households are drawn with systematic sampling from the General Registry Office lists of each municipality.

Each year the survey collects more than 60 detailed categories of expenditure; following Perali (1997), we have aggregated them in seven broader consumer good groups:

- food;
- alcohol and tobacco;
- clothing;
- fuel and heating;
- transport;
- housing, health and education;
- entertainment and other.

3.2 Nonparametric regression Engel curves

A preliminary analysis has been conducted in order to obtain an immediate picture of the aggregate information; it has been done with the ultimate goal of investigating empirically the Engel curve shapes.

A kernel nonparametric regression estimator has been used to analyse the relationship between the shape of each of the seven aggregated categories of consumer goods, and the logarithm of the total consumption expenditure:

$$\hat{m}(x) = \frac{1}{n} \sum W(x) y_i \tag{21}$$

where $W(x)$ is a weighting system depending on the kernel function K(.) and on the bandwidth parameter h. In nonparametric regression we have utilised a normal kernel function, and we have evaluated the parameter h with a CROSS-VALIDATION procedure (Hardle 1990, Betti 1999).

Shares of the seven categories have been regressed on the logarithm of the total expenditure; the Appendix reports the seven curves estimated for the year 1994; corresponding curves for previous years give similar shapes.

Pictures A1 and A7 show that only Engel curves for food or entertainment and other present approximating linear shape. All the other curves are clearly nonlinear; in particular curves corresponding to clothing or fuel and heating exhibit a quadratic shape.

3.3 The price system and the demand rank

The BF survey reports information on expenditure on goods; it is therefore impossible to disaggregate into quantities and prices. These variables are necessary in the estimation of any complete demand system. In fact, from detailed expenditures, it is possible to analyse the relations between shares and total expenditure, but elasticities with respect to prices can be estimated only by imposing a priori some strong assumptions; such assumptions are often rejected by empirical evidence.

The prices households face have two directions of variability. Prices change exogenously with respect to time and space; and a second spatial and endogenous variability is intrinsic in the consumer's choice.

In this paper we reconstruct the implicit price indices following Lewbel (1989,b); the index for time t (with base time 0), for the commodity goods group i, faced by the household with characteristics $\mathbf{z}$ and consumption $\mathbf{q}$, is defined as follow:

$$ {}_0J_t^i\left(\mathbf{q},\mathbf{z}\right) = {}_0I_t^i \times M_i\left(\mathbf{q},\mathbf{z}\right) \tag{22} $$

where ${}_0I_t^i$ is the official price index for goods group i at time t (with base 0), published by the Italian Office of Statistics; while $M_i\left(\mathbf{q},\mathbf{z}\right)$ is the index which explains the variability intrinsic in the consumer choice. The methodology here utilised to calculate $M_i\left(\mathbf{q},\mathbf{z}\right)$ is not reported in the paper; interested readers can consult Lewbel (1989,b) or Betti (1998). Table 3.1 presents the average value of ${}_0J_t^i\left(\mathbf{q},\mathbf{z}\right)$ over all households for each group and year.

Table 3.1: The price system

Year Group	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Food	0.923	0.969	1.010	1.044	1.111	1.180	1.259	1.321	1.345	1.391
Alcoholic and tobacco	0.705	0.875	0.857	0.912	0.981	1.109	1.203	1.289	1.293	1.321
Clothing	0.817	0.878	0.975	1.034	1.093	1.154	1.213	1.270	1.297	1.332
Fuel and heating	0.848	0.845	0.880	0.919	0.980	1.074	1.155	1.197	1.258	1.330
Transport	0.784	0.817	0.843	0.884	0.928	0.986	1.044	1.098	1.160	1.238
Housing, health and education	1.184	1.352	1.407	1.535	1.650	1.880	2.011	2.310	2.370	2.420
Entertainment and others	0.851	0.919	0.968	1.026	1.108	1.185	1.287	1.400	1.499	1.579

Once the prices have been calculated, it is possible to apply a nonparametric test for nonlinearities in the Engel curves: the rank test. The rank of a demand systems is defined as the maximum dimension spanned by the Engel curves.

Any demand system in budget shares $\mathbf{w} = (w_1, w_2, ..., w_k)$, can be defined as:

$$w_i = \sum_{r:1}^{R} A_{r,i}(\mathbf{p})\, G_{r,i}(y, \mathbf{p}) \tag{23}$$

where $y = \log C(u, \mathbf{p})$, $A(\mathbf{p})$ is a R-rank matrix for some price system $\mathbf{p}$, and $G_i(y, \mathbf{p})$ is an R x 1 vector. The rank of $\mathbf{w}$ is the minimum dimension R so that equation (23) can be satisfies for any $\mathbf{p}$ and y.

Gill and Lewbel (1992) proposes a nonparametric test for the rank of the demand system; this consists in the estimation of the pivots of the matrix representing the association between expenditure and shares.

Table 3.2: Estimating pivots and LDU rank test, 1985 - 1994.

Pivots		Test	d.of f.	X^2	p-value
d_1	0.6880	R = 1	4	36600	0.000
d_2	0.1213	R = 2	3	719	0.000
d_3	0.0074	R = 3	2	0.360	0.996
d_4	0.0017	R = 4	1	0.075	0.998

Under the null hypothesis all the pivots, for r larger than R, are zero; therefore the test is against the alternative hypothesis that the rank is greater than R. Table 3.2 reports the value for four tests based on the pooled data from 1985 to 1994; the null hypothesis is rejected in the first two tests and not rejected in the third: the rank of the demand system is therefore taken to be three.

4 Empirical analysis

The analysis has been conducted in two directions, with the aim of selecting the best fitting complete demand system and the best set of demographic variables. All the models have been estimated using pooled data from 1985 to 1994, and program TSP version 4.3.

4.1 The complete demand systems

The first simple estimated model is the AI system of Deaton and Muellbauer (1980):

$$w_i = \alpha_i + \sum_{j:1}^{7} \gamma_{ij} \ln p_j + \beta_i \ln x. \tag{24}$$

In this model the parameter α_0 (contained in lnx) is not identified (Pashardes, 1989);

but for a consumer goods group whose parameter α_i was close to zero (i.e alcohol and tobacco), it would be possible to estimate α_0.

A preliminary analysis, conducted imposing some parameters to be equal to zero, has found the optimal value for $\widehat{\alpha}_0$ to be equal to 14.06. Moreover, since the TSP program estimates the model parameters with the iterative procedures *lsq* (least squares), the AIDS estimates are utilised as starting values for the following models.

Column (1) in Table 4.1 reports the AIDS model estimates. All the α_i and β_i parameters differ significantly from zero; α_7 and β_7 can be calculated by subtraction since:

$$\sum \alpha_i = 1 \Rightarrow \alpha_7 = 1 - \sum_{i:1}^{6} \alpha_i \tag{25}$$

$$\sum \beta_i = 0 \Rightarrow \beta_7 = -\sum_{i:1}^{6} \beta_i . \tag{26}$$

The next step consists in introducing the quadratic term into the consumption shares:

$$w_i = \alpha_i + \sum_{j:1}^{7} \gamma_{ij} \ln p_j + \beta_i \ln x + \frac{\phi_i}{b(\mathbf{p})} (\ln x)^2 \tag{27}$$

and in the introduction of the σ-concave CES transformation in the price kernel; this leads to the CES_QUADIS budget shares of the following form:

$$w_i = \frac{\alpha_i p_i^{(1-\sigma)}}{\sum \alpha_k \ln p_k^{(1-\sigma)}} + \sum \gamma_{ij} \ln p_j + \beta_i \ln x + \frac{\phi_i}{b(\mathbf{p})} (\ln x)^2 . \tag{28}$$

Column (2) in Table 4.1 reports the CES_QUAIDS model estimates. The large t-values are evidence that the ϕ_i parameters and σ are different from zero; this analysis confirms the high nonlinearities in the Engel curves, already discovered in the preliminary analysis in section 3. The parameter ϕ_7 can be calculated from the homogeneity condition:

$$\sum \phi_i = 0 \Rightarrow \phi_7 = -\sum_{i:1}^{6} \phi_i . \tag{29}$$

The model choice is completed with the introduction of demographic variables using the *Price Scaling* model proposed by Ray (1983), whose corresponding cost function in logarithm form is the following:

$$\ln C\,(u,\mathbf{p},\mathbf{z}) = \ln m(\mathbf{p},\mathbf{z}) + \ln C_R(u,\mathbf{p}) \tag{30}$$

and substituting into the CES_QUAIDS cost function, leads to the CES_QUAIDS_PS:

$$\ln C\,(u,\mathbf{p},\mathbf{z}) = \ln a(\mathbf{p},\mathbf{z}) + \frac{ub(\mathbf{p})}{1-u-\phi(\mathbf{p})} \tag{31}$$

where

$$\ln a\,(\mathbf{p},\mathbf{z}) = \ln m_0(\mathbf{p},\mathbf{z}) + \ln a(\mathbf{p}). \tag{32}$$

Three different specifications are chosen for $\ln m_0(\mathbf{p},\mathbf{z})$. In the first one, the only demographic variable introduced is the household size z_s:

$$\ln m_0(\mathbf{p},z_s) = \ln m_0\,(z_s) = \ln\left(1+\tau_0 z_s\right). \tag{33}$$

The estimated model is reported in column (3) in Table 4.1. The value $\widehat{\tau}_0 = 0.176$, although significantly different from zero, leads to economies of scale too large; consequently the corresponding equivalence scale would be too flat.

In the second specification four demographic variables are considered:

z_0 = number of adults over 18;

z_1 = number of children aged 0 - 5;

z_2 = number of children aged 6 - 14;

z_3 = number of children aged 15 - 18:

$$\ln m_0(\mathbf{p},\mathbf{z}) = \ln m_0\,(\mathbf{z}) = \ln\left(1+\sum_{i=0}^{3}\tau_i z_i\right). \tag{34}$$

Table 4.1: Estimating five complete demand systems*

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (5)	
α_1	0.3599	0.3981	0.4007	0.4014	0.3260		
	(88.8529)	(24.5197)	(12.5580)	(12.5938)	(6.8582)		
α_2	0.0763	0.0543	0.0790	0.0789	0.1003		
	(24.7244)	(5.5818)	(6.2867)	(6.2779)	(10.5012)		
α_3	0.0501	0.0622	0.0541	0.0543	0.0429		
	(20.4877)	(11.3051)	(8.4346)	(8.4547)	(5.7725)		
α_4	0.0920	0.0507	0.0779	0.0776	0.1158		
	(22.5244)	(4.0195)	(4.1621)	(4.1685)	(2.7191)		
α_5	0.0451	0.0437	0.0493	0.0494	0.0441		
	(19.5082)	(15.6579)	(20.9348)	(20.8833)	(10.0947)		
α_6	0.3435	0.3515	0.2806	0.2806	0.2889		
	(98.7361)	(27.5592)	(29.1524)	(29.6155)	(36.4724)		
β_1	-0.9059	-0.0644	-0.0973	-0.0979	-0.1077		
	(-1.6972)	(-7.8238)	(-12.8224)	(-13.0141)	(-17.5220)		
β_2	0.0241	0.0426	0.0402	0.0405	0.0216		
	(12.2816)	(7.0171)	(4.6188)	(4.6922)	(2.8353)		
β_3	-0.0174	-0.0232	-0.0193	-0.0195	-0.0158		
	(-13.4528)	(-7.7925)	(-7.5181)	(-7.5376)	(-11.6346)		
β_4	0.0730	0.0517	0.0571	0.0572	0.0973		
	(20.0159)	(6.5570)	(4.4708)	(4.4499)	(2.3462)		
β_5	-0.0053	0.0044	-0.0021	-0.0022	-0.0085		
	(-5.6574)	(1.3031)	(-0.6107)	(-0.6347)	(-4.0212)		
β_6	-0.0169	-0.0522	-0.0259	-0.0256	-0.0131		
	(-4.5764)	(-4.9946)	(-2.3125)	(-2.2863)	(-0.6897)		
ϕ_1		-0.0130	-0.0095	-0.0095	-0.0065		
		(-5.2767)	(-3.0943)	(-3.0506)	(-2.5050)		
ϕ_2		-0.0101	-0.0132	-0.0132	-0.0082		
		(-4.2366)	(-4.5185)	(-4.5790)	(-3.6830)		
ϕ_3		0.0029	0.0025	0.0025	0.0006		
		(2.6953)	(2.0393)	(2.0923)	(0.5855)		
ϕ_4		0.0133	0.0198	0.0201	0.0480		
		(3.3499)	(4.0869)	(4.0914)	(10.4302)		
ϕ_5		-0.0050	-0.0044	-0.0048	-0.0023		
		(-5.1256)	(-3.5728)	(-3.5968)	(-2.1093)		
ϕ_6		0.0183	0.0166	0.0167	-0.0218	-0.2470	δ_1
		(4.8958)	(3.7449)	(3.7242)	(-6.0674)	(-5.3612)	
σ		1.8700	1.6904	1.6875	1.6692	-0.0776	δ_2
		(16.1074)	(14.0440)	(14.0181)	(14.4681)	(-2.3337)	
τ_0			0.1764	0.2281	0.2564	0.0244	δ_3
			(8.5825)	(7.3706)	(3.5301)	(0.9382)	
τ_1				0.0839	0.0999	0.4016	δ_4
				(2.1096)	(1.6138)	(10.9525)	
τ_2				0.1746	0.1845	0.1464	δ_5
				(4.9399)	(2.6068)	(4.1103)	
τ_3				0.1453	0.2471	-0.4753	δ_6
				(3.0752)	(2.7966)	(-14.6960)	

* t-statistic values are reported in parentheses.

Column (4) in Table 4.1 reports parameter estimates for model (34). The estimated coefficient $\widehat{\tau}_2 = 0.175$ corresponding to children aged 6 - 14 is larger than $\widehat{\tau}_3 = 0.145$ corresponding to children aged 15 - 18. This could imply that child costs in the 6 - 14 group are higher than those in the 15 - 18 group. This misleading result is probably due to the lack of interaction between prices and demographic variables in the specification of $\ln m_0(\mathbf{p}, \mathbf{z})$. We consider now the model containing the interaction between prices and demographic variable as follows:

$$\ln m_0(\mathbf{p}, \mathbf{z}) = \ln \left(1 + \sum_{i=0}^{3} \tau_i z_i\right) + \sum z_s \delta_k p_k \tag{35}$$

where z_s is the household size.

From column (5) in Table 4.1 it is possible to note that the interaction parameters δ_i are significantly different from zero, except δ_3, whose t-statistic value is equal to 0.938. Including this interaction leads to progressively increasing parameter estimates $\widehat{\tau}_1$, $\widehat{\tau}_2$ and $\widehat{\tau}_3$, in line with the age group. In an unreported further model we have introduced a new set of demographic characteristics: we have distinguished between adults and elderly people. The corresponding estimated parameters are almost equal, showing that this modification is unnecessary.

4.2 Equivalence scales

Equivalence scales are calculated with the parameters reported in column (5) in Table 4.1, and corresponding to the model CES_QUAIDS_PS. From the definition (1):

$$e(u, \mathbf{p}, \mathbf{z}, \mathbf{z}^r) = \frac{C(u, \mathbf{p}, \mathbf{z})}{C(u, \mathbf{p}, \mathbf{z}^r)} = \frac{a(\mathbf{p}, \mathbf{z}) \exp\left\{\frac{ub(\mathbf{p})}{1-u-\phi(\mathbf{p})}\right\}}{a(\mathbf{p}, \mathbf{z}^r) \exp\left\{\frac{ub(\mathbf{p})}{1-u-\phi(\mathbf{p})}\right\}} \tag{36}$$

and substituting $a(\mathbf{p}, \mathbf{z})$ from (32) into (36):

$$e(u, \mathbf{p}, \mathbf{z}, \mathbf{z}^r) = \frac{m_0(\mathbf{p}, \mathbf{z}) a(\mathbf{p})}{m_0(\mathbf{p}, \mathbf{z}^r) a(\mathbf{p})} = \frac{m_0(\mathbf{p}, \mathbf{z})}{m_0(\mathbf{p}, \mathbf{z}^r)} \tag{37}$$

one gets the equivalence scale reported in column (5) in Table 4.2.

The Table reports three other equivalence scales: i) the OECD 50-30 scale, here converted to let the couple be the reference household; ii) nonparametric equivalence scales calculated by Betti (1998) using the 1993 Bank of Italy Survey on Italian Household Budget; iii) the Italian official scale proposed by the Commissione Gorrieri (1986),

based on a food ratio linear model for years 1982-1984 (as in Van Ginneken, 1982).

The CES_QUAIDS scale clearly presents economies of scale higher than the others; this fact is particularly interesting in the comparison with the OECD scale. In fact, since nonparametric and Gorrieri scales are both based on the food ratio method, they constitute an overestimate of the scale compared with those calculated with a complete demand system (Deaton and Muellbauer, 1986).

Table 4.2: CES_QUAIDS_PS equivalence scale

Adult	Ch 0 - 5	Ch 6 - 14	Ch 15 - 18	Scale	OECD	N.P.	Gorrieri
1	0	0	0	0.81	0.67	0.59	0.60
1	1	0	0	0.85	0.87	1.00	1.00
1	0	1	0	0.95	0.87	1.00	1.00
1	0	0	1	1.02	0.87	1.00	1.00
2	0	0	0	1.00	1.00	1.00	1.00
1	1	1	0	1.14	1.07	1.24	1.33
1	1	0	1	1.17	1.07	1.24	1.33
1	0	1	1	1.28	1.07	1.24	1.33
2	1	0	0	1.07	1.20	1.24	1.33
2	0	1	0	1.14	1.20	1.24	1.33
2	0	0	1	1.17	1.20	1.24	1.33
3	0	0	0	1.22	1.33	1.24	1.33
2	1	1	0	1.28	1.40	1.42	1.63
2	0	1	1	1.32	1.40	1.42	1.63
2	1	0	1	1.35	1.40	1.42	1.63
3	1	0	0	1.19	1.53	1.42	1.63
3	0	1	0	1.32	1.53	1.42	1.63
3	0	0	1	1.49	1.53	1.42	1.63
4	0	0	0	1.41	1.67	1.42	1.63
4	1	0	0	1.33	1.87	1.72	1.90
4	0	1	0	1.62	1.87	1.72	1.90
4	0	0	1	1.69	1.87	1.72	1.90
5	0	0	0	1.63	2.00	1.72	1.90
6	0	0	0	1.93	2.33	1.96	2.16

5 Concluding remarks

In this paper we have estimated equivalence scales based on complete demand systems.

A preliminary analysis, conducted by means of nonparametric regression, has shown some nonlinearities in the Engel curves for seven categories of consumer goods, from the Italian Household Budget Survey 1985 - 1994. This analysis has been confirmed by testing for the rank using the pooled data set.

Starting from the Almost Ideal Demand System, we have introduced nonlinearities in the underlying Engel curves, through a quadratic component. Moreover several sets

of socio-demographic characteristics have been introduced by means of the *Price Scaling* method.

Empirical evidence shows that the quadratic specification of the AIDS model fits the pooled data very well; moreover the best set of demographic characteristics has been identified as the number of adults and the numbers of children in three different age categories. The estimated model leads to equivalence scales presenting strong economies of scale, compared with scales based on food ratio, and with the OECD 50 - 30 scale.

The set of demographic variables considered does not distinguish between adults and elderly people; empirical analysis has shown that people in the age category 65+ have the same weight as adults in the age category 18 - 64. This fact is probably due to the peculiar structure of the Italian households, where young adults still live together at home with elderly parents.

Possibly the model could be developed to differentiate households which are all elderly from those with both young and elderly adults.

Acknowledgement 2 *I would like to thank Vincenzo Atella, Raffaella Castagnini, Federico Perali, Elena Quercioli, Nicola Rossi and particularly Jane Galbraith for helpful comments.*

6 Appendix: Engel curves for seven commodity categories, fitted using nonparametric regression.

Note: in each of the seven Pictures A1 - A7, the vertical axis is the proportion of expenditure that is spent on that commodity, and the horizontal axis is the logarithm of total expenditure.

References

[1] Banks J., Blundell R. and Lewbel A. (1997), "Quadratic Engel Curves and Consumer Demand", *The Review of Economics and Statistics* **79**, pp. 527-539.

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

[3] Betti G. (1998), "Scale di equivalenza intertemporali: un'applicazione sulla base dello pseudo panel dell'Indagine ISTAT sui Consumi delle Famiglie 1985-1994", Ph.D. Thesis, University of Florence.

[4] Betti G. (1999), "Nonparametric Equivalence Scales with Application to Poland", *Statistics Research Report* LSERR49, London School of Economics.

[5] Blackorby C. and Donaldson D. (1989), "Adult-equivalence scales, interpersonal comparisons of well-being, and applied welfare economics", Discussion paper n. 89/24, University of British Columbia, Vancouver.

[6] Bollino C.A. and Rossi N. (1989), "Demographic Variables in Demand Systems and Related Measures of the Cost of Changing Family Size", *Giornale degli Economisti e Annali di Economia* **48**, pp. 449-465.

[7] Buhman B., Rainwater G., Schmaus G. and Smeeding T.M. (1988), "Equivalence scales, well-being, inequality and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database", *Review of Income and Wealth* **34**, pp. 115-142.

[8] Commissione Gorrieri di Indagine sulla Povertà, (1986), "Primo Rapporto e Studi di Base", Presidenza del Consiglio dei Ministri, Roma.

[9] Deaton A.S. and Muellbauer J. (1980), "An Almost Ideal Demand System", *Americn Economic Review* **70**, pp. 312-326.

[10] Deaton A.S. and Muellbauer J. (1986), "On Measuring Child Cost: With Applications to Poor Countries", *Journal of Political Economy* **94**, pp. 720-744.

[11] Gill L. and Lewbel A. (1992), "Testing the Rank and Definiteness of Estimated Matrices with Applications to Factor, State-space and Arma Models", *Journal of the American Statistical Association* **87**, pp. 766-776.

[12] Gorman W. (1976), "Tricks with Utility Functions", in M.J. Artis and A.R. Nobay (eds.) "*Essays in Economic Analysis*", Proceedings of the 1975 AUTE Conference, Sheffield. Cambridge: Cambridge University Press.

[13] Hagenaars A.J.M., de Vos K. and Zaidi M.A. (1994), "Poverty Statistics in the Late 1980s: Research based on micro-data", EUROSTAT, Luxembourg.

[14] Hardle W. (1990), "*Applied Nonparametric Regression*". Cambridge University Press.

[15] Lancaster G. and Ray R. (1998), "Comparison of Alternative Models of Household Equivalence Scales: The Australian Evidence on Unit Record data", *Economic Record* **74**, pp. 1-14.

[16] Lewbel A. (1985), "A Unifying Approach to Incorporating Demographic or Other Effects into Demand Analysis", *Review of Economic Studies* **52**, pp. 1-18.

[17] Lewbel A. (1989,a), "Household Equivalence Scales and Welfare Comparisons", *Journal of Public Economics* **39**, pp. 377-391.

[18] Lewbel A. (1989,b), "Identification and estimation of equivalence scales under weak separability", *Review of Economic Studies* **56**, pp. 311-316.

[19] Lewbel A. (1991), "The rank of demand systems: theory and non-parametric estimation", *Econometrica* **59**, pp. 711-730.

[20] Muellbauer (1976), "Community preferences and the representative consumer", *Econometrica* **44**, pp. 979-999.

[21] Pashardes P. (1989), "Contemporaneous and intertemporal child costs: Equivalent expenditure vs. equivalent income scales", Working paper n. 89/2, Institute for Fiscal Studies, London.

[22] Perali F. (1997), "Stima delle scale di equivalenza con i dati elementari dell'indagine sui Bilanci Familiari ISTAT 1985-1994", mimeo, University of Verona.

[23] Pollak R.A. and Wales T.J. (1978), "Estimation of Complete Demand Systems from Household Budget Data: the Linear and Quadratic Expenditure System", *American Economic Review* **69**, pp. 216-221.

[24] Pollak R.A. and Wales T.J. (1981), "Demographic Variables in Demand Analysis", *Econometrica* **49**, pp. 1533-1551.

[25] Ray R. (1983) "Measuring the Cost of Children: An Alternative Approach", *Journal of Public Economics* **22**, pp. 89-102.

[26] Van Ginneken W. (1982), "Generating Internationally Comparable Income Distribution Data: Evidence from the Federal Republic of Germany (1974), Mexico (1968) and the United Kingdom (1979)", *The Review of Income and Wealth* **28**, pp. 365-379.