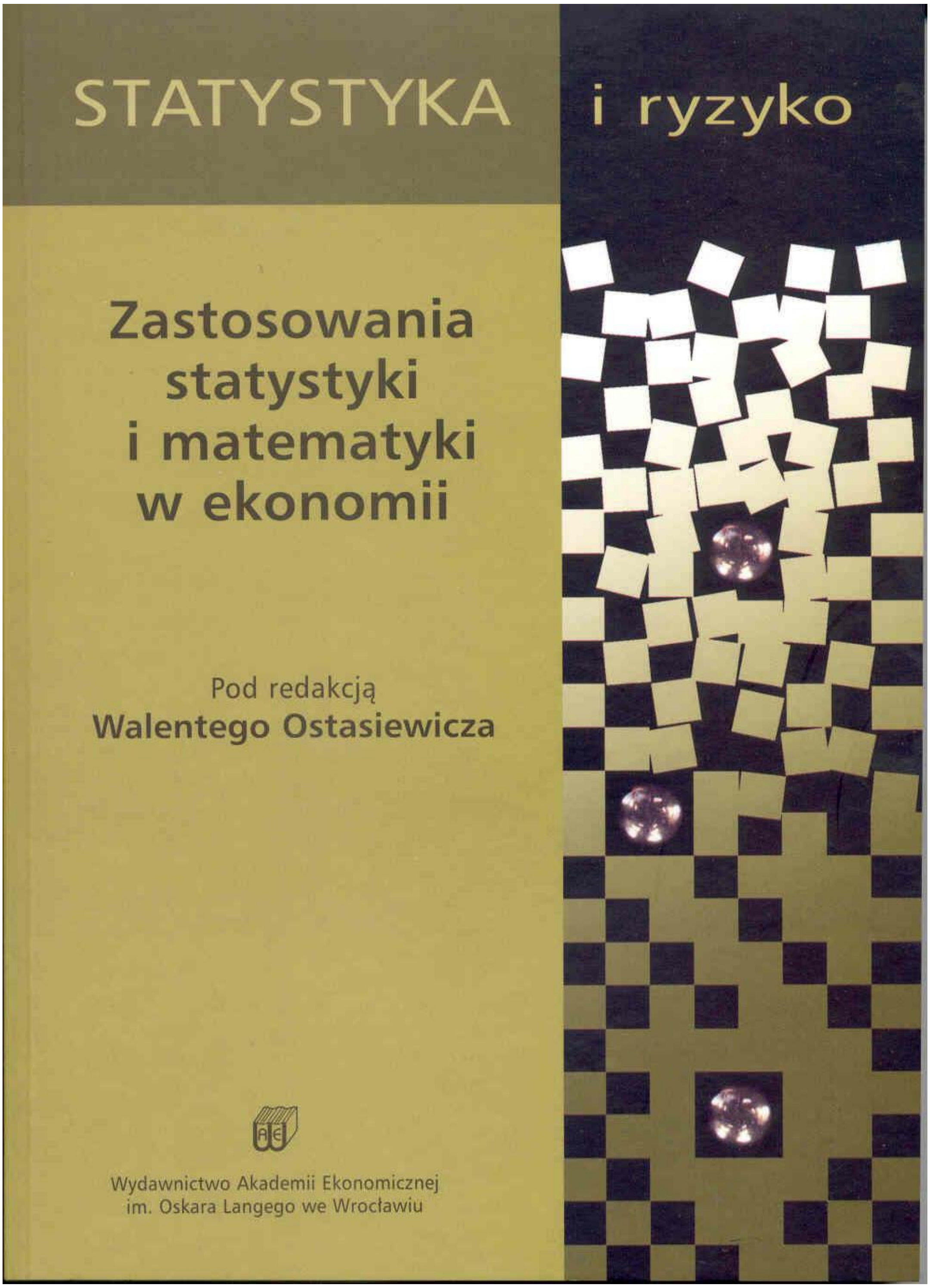

STATISTICAL AND MATHEMATICAL APPLICATIONS IN ECONOMICS

EDITED BY WALENTY OSTASIEWICZ

SERIES STATISTICS AND RISKS

UNIVERSITY OF ECONOMICS OSKAR LANGE
WROCLAW 2004
PL ISSN 0324-8445

PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1036 2004
Zastosowania statystyki i matematyki w ekonomii

Achille Vernizzi, Elena Siletti
Università degli Studi di Milano

ESTIMATING THE COST OF CHILDREN THROUGH ENGEL CURVES BY DIFFERENT GOOD AGGREGATES

1. Introduction

The original idea underlying Engel's method is that the standard of living of a household is negatively related to the share of the household budget spent on food. Engel derived his idea on the observation of empirical regularities from which he formulated his famous laws: (1) the share of budget devoted to food decreases as income increases and (2) increases, ceteris paribus, as the number of family components increases [3].

Notwithstanding more sophisticated models has been developed and proposed in literature on the basis of household behaviour theory, the Engel approach has been widely applied to estimate equivalence scales even recently: for example, in Italy the means test for the access to public services and related fares are evaluated by a scale based on estimates obtained through this approach (see [1]).

Even if many authors are of the opinion that Engel curves can still be a valid investigation tool, some of them suggest (see [4]) that the share adopted as welfare indicator should not be restricted to food, but should include a wider number of goods. They argue that in a modern economy food cannot be any longer an exclusive measure of welfare either because people tend to consume meals outside home much more often than they were used a hundred years ago or because the essential needs in an advanced society include items outside the category of food.

For instance, when investigating on the cost of children, there are other expenses in addition to food that families have to face for their children, which are considered rather as necessary than as unnecessary: for example this the case for education, transports, school books, school fees and so on.

In this article we shall estimate equivalence scales for the cost of children in one of the most developed Italian region: Lombardia. The data set is obtained from the survey on household consumption expenditures, which is hold by the Italian Central Institute for Statistics (ISTAT). In this paper we consider the data collected in Lombardia in 1997, 1998, 1999 and 2000: it must be stressed that ISTAT's survey is not structured as a panel, but each year families to be interviewed are randomly drawn through a two stage sampling.

2. The models

According to most part of literature, first of all, we specify Engel's laws in models where the share demanded for food, with respect to total expenditure, is expressed as a function of the logarithm of total expenditure itself, of dummies conceived to take account of the different number of children and their age and, moreover, of other variables specified to keep account of the main heterogeneities among households.

Moreover, as it is shown afterwards, we introduce as dependent variable, that is as welfare indicator, four further different shares, obtained by aggregating other goods together with food and then dividing the value of the resulting aggregate by the value of total expenditure.

In most literature these models are formalised specifying the dependent variable directly in the form of a share; in a rather simplified version, with family characteristics restricted to the number of components, the typical equation to be estimated is:

$$w_h = \alpha + \beta \log y_h + \eta g\left(n_h\right) + \varepsilon_h \qquad (1)$$

where:

- the subscript h refers to the h-th household (h=1,2,...,H);
- w_h is the aggregate share (food) observed for household h;
- y_h is total expenditure observed for the h-th household, y_h is expressed in nominal prices;
- $g(n_h)$ is a suitable function of the number of components in the h-th household;

- ε_h is an unknown random component for which standard hypothesis are introduced;
- α, β, η are unknown coefficients to be estimated.

However Carbonaro [1985] adopts a model, which takes the following form:

$$\log q_h = \alpha + \gamma \log y_h + \eta g(n_h) + \varepsilon_h, \tag{2}$$

- q_h is the quantity (value) demanded for food by household *h*.

We easily observe that model (2) can be written in the following equivalent way:

$$\log q_h - \log y_h = \alpha + \gamma \log y_h - \log y_h + \eta g(n_h) + \varepsilon_h$$

that is:

$$\log \frac{q_h}{y_h} = \alpha + (\gamma - 1)\log y_h + \eta g(n_h) + \varepsilon_h$$

or, in terms of share and defining $(\gamma - 1) = \beta$:

$$\log w_h = \alpha + \beta \log y_h + \eta g(n_h) + \varepsilon_h. \tag{2'}$$

In order to avoid any priori choice between (1) and (2), we decide to adopt the Box-Cox transformation for the dependent variable; in fact, as it is well known, this transformation nests both (1) and (2):

$$\frac{w_h^\lambda - 1}{\lambda} = \alpha + \beta \log y_h + \eta g(n_h) + \varepsilon_h. \tag{3}$$

The boundary cases for (3) are λ=1, when (3) reduces to (1) and, $\lim_{\lambda \to 0} \frac{w^\lambda - 1}{\lambda} = \log w$, when (3) takes the form of (2').

In our case the typical equation to be estimated becomes then.

$$\frac{{}_g w_h^\lambda - 1}{\lambda} = \alpha + \beta \log y_h + \boldsymbol{\theta}'\mathbf{d}_h + \boldsymbol{\psi}'\mathbf{z}_h + \tau_1 t_{97} + \tau_2 t_{98} + \tau_3 t_{99} + \varepsilon_h. \tag{4}$$

- $\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{\psi}$ is the set of unknown coefficients to be estimated;

- $\mathbf{d}_h$ is a vector which contains suitable dummies introduced to take account of the age and of the number of children in the h-th household: as we show afterwards, the ten different cases are taken into consideration;
- $\mathbf{z}_h$ is a vector of variables introduced to control for family characteristics we are not directly interested to, which, if omitted, might give biased estimates for β, $\boldsymbol{\theta}$ and α;
- t_{97}, t_{98} and t_{99} are three dummies, one for each year: they have been introduced to capture the effect of the level of prices and others unspecified characteristics of the year they refer to.

In (4) ${}_g w_h$, the aggregate share, is now indexed by a g due to the fact that the shares used as welfare index include good sets that can be broader than food. As it can be seen afterwards, we try five different aggregates, so g=1,2,3,4,5.

Equation (4) implies that, given $\mathbf{d}_h$ and $\mathbf{z}_h$, the demanded share for food (after the Box-Cox transformation) is linear w.r.t. the logarithm of income (total expenditure); given the level of income and $\mathbf{z}_h$, the demand is shifted according to the value assumed by $\mathbf{d}_h$.

If two families, R and h, show the same share for the demand of aggregate g, assuming that $\mathbf{z}_h$ is the same for both families, we easily get that the ratio of the two incomes is:

$$\frac{y_h}{y_R} = \exp\left\{-\frac{1}{\beta}\left[\boldsymbol{\theta}'\mathbf{d}_h - \boldsymbol{\theta}'\mathbf{d}_R\right]\right\} = s(\mathbf{d}_h). \quad (5)$$

If the share ${}_g w$ is a reliable indicator of the standard of living, family h to enjoy the same welfare as family R, should have the income of the latter, y_R, times $s(\mathbf{d}_h)$, which is the equivalence scale of family h w.r.t family R. The scale (5) does not depend on the level of utility enjoyed nor on other variables, but on the number of children and their characteristics specified through vector $\mathbf{d}$.

3. Household characteristics

The data we use in this analysis are from the *Indagine sui Consumi delle Famiglie Italiane*[1], ISTAT[2] The survey population is the whole set of

[1] Italian Household Expenditure Survey.
[2] ISTAT: Italian Central Institute for Statistics.

households resident in Italy. We focus our study only on households resident in Lombardia[3].

Every year the survey concerns 24.000 households in Italy, and more then 2000 in Lombardia.

The surveys considered are those performed in 1997, 1998, 1999 and 2000.

Table 1. Total number of households per year in Lombardia

	1997	1998	1999	2000
N° Households	2 330	2 160	2 068	2 463

The aim of our research is to estimate the cost of children, so initially we limit our dataset to households formed by a couple of adults without children, a couple with one child, with two, three, and four children and where the reference person is from 20 to 60 years old.

The boundaries of the age range are chosen to select just couples of potential parents.

Table 2. Number of households in the restricted dataset

	1997	1998	1999	2000	Total
Total	1 304	1 194	1 060	1 234	4 792
Couples	253	267	253	290	1 063
Couples with 1 child	449	434	362	436	1 681
Couples with 2 children	495	402	364	414	1 675
Couples with 3 children	94	77	72	85	328
Couples with 4 children	13	14	9	9	45

From Table 3 and Table 4 we observe that both the percentage of married couples and that of housewives show a relation with the number of children. This relation is quite evident for the percentage of housewives: it clearly increases up to three children, and it drops for couples with four children. A similar trend, even if less remarked, can be found in Table 3. Before trying to interpret the lowering of the percentages observed for families with four children, we should remember that the number of these families in the sample is rather low: they are only 45 in the whole period.

[3] Lombardia is a northern region of Italy.

Table 3. Percentage of married couples w.r.t. the number of children

	Married Couples
Couples	92.1%
Couples with 1 child	97.3%
Couples with 2 children	98.9%
Couples with 3 children	98.9%
Couples with 4 children	96.7%

Table 4. Percentage of Housewives w.r.t. the number of children

Women	Housewives	Workers
Couples	24.4%	63.0%
Couples with 1 child	30.1%	59.7%
Couples with 2 children	39.7%	51.4%
Couples with 3 children	52.3%	40.8%
Couples with 4 children	30.0%	63.3%

Reference persons show dissimilar school levels in different household typologies. Table 5 shows that the share of post-graduated and graduated is highest for couples with four children and for couples without children; in the meanwhile, it is worth to stress, that the former typology shows the highest percentage of low-level people.

Not surprisingly, the average age of reference persons increases with the number of children; this is due to the selection operated on the sample.

Table 5. School levels of the household reference persons

Reference Person	Post-Graduated, Graduated	Medium level	Low level
Couples	13.0%	38.3%	48.6%
Couples with 1 child	10.8%	35.3%	53.9%
Couples with 2 children	11.1%	35.7%	53.2%
Couples with 3 children	11.5%	36.7%	52.8%
Couples with 4 children	13.3%	16.6%	70.1%

Table 6. Average ages of the household reference persons

Reference person	Average age
Couples	43.2
Couples with 1 child	45.0
Couples with 2 children	45.8
Couples with 3 children	47.2
Couples with 4 children	49.1

Table 7. Households with at most 2 children: age differentiations

Household typology	Year	1997	1998	1999	2000	Total
Total		858	772	673	782	3085
Couples with 1 child		395	386	323	393	1497
Couples with 1 child <3		58	45	37	53	193
Couples with 1 child (3-5)		37	59	32	51	179
Couples with 1 child (6-13)		97	88	79	87	351
Couples with 1 child student (14-19)		49	49	37	51	186
Couples with 1 child student >20		95	91	78	90	354
Couples with 1 child out of ranges		59	54	60	61	234
Couples with 2 children		463	386	350	389	1588
Couples with 2 children <3		0	0	4	2	6
Couples with 1 child <3 & 1 child (3-5)		34	23	11	25	93
Couples with 1 child <3 & 1 child (6-13)		23	15	19	18	75
Couples with 1 child <3 & 1 child student (14-19)		2	0	1	1	4
Couples with 1 child <3 & 1 child student >20		0	0	0	0	0
Couples with 2 children (3-5)		4	4	10	5	23
Couples with 1 child (3-5) & 1 child (6-13)		50	36	32	41	159
Couples with 1 child (3-5) & 1 child student (14-19)		5	4	1	2	12
Couples with 1 child (3-5) & 1 child student >20		0	1	0	0	1
Couples with 2 children (6-13)		60	34	41	57	192
Couples with 1 child (6-13) & 1 child student (14-19)		54	45	41	49	189
Couples with 1 child (6-13) & 1 child student >20		3	7	7	6	23
Couples with 2 children students (14-19)		14	19	15	10	58
Couples with 1 child student (14-19) & 1 child student >20		17	15	12	19	63
Couples with 2 children students >20		11	8	2	6	27
Couples with 2 children out of ranges		186	175	154	148	663

In order to introduce tractable children age differentiations among households, we contemplate only households with at most two children. On the other hand, families with more than two children are less than the 8% of the total number of families in the sample.

To deal just with children supported by parents, we discriminate households where children from 14 to 19, and over 20 years are students from households where children, in these age ranges, are not students, because we consider the former groups as economically dependent from parents; as the latter groups may include both economically dependent and not dependent children, we do drops them from our analysis and concentrate ourselves just on students. The whole cases considered and their frequencies are reported in Table 7.

The classification adopted to estimate equivalence scales is a simplification of Table 7. We use ten groups: the first five groups with households with one child:

- 1st Group: couples with 1 child less than three years old
- 2nd Group: couples with 1 child between 3 and 5 years old
- 3rd Group: couples with 1 child between 6 and 13 years old
- 4th Group: couples with 1 child between 14 and 19 years old who is a student
- 5th Group: couples with 1 child more than twenty years old who is a student

Table 8. Aggregation groups for couples with two children

6 th Group	59	38	35	46	178
Couples with 2 children <3	0	0	4	2	6
Couples with 1 child <3 & 1 child (3-5)	34	23	11	25	93
Couples with 1 child <3 & 1 child (6-13)	23	15	19	18	75
Couples with 1 child <3 & 1 child student (14-19)	2	0	1	1	4
7 th Group	59	45	43	48	195
Couples with 2 children (3-5)	4	4	10	5	23
Couples with 1 child (3-5) & 1 child (6-13)	50	36	32	41	159
Couples with 1 child (3-5) & 1 child student (14-19)	5	4	1	2	12
Couples with 1 child (3-5) & 1 child student >20	0	1	0	0	1
8 th Group	60	34	41	57	192
Couples with 2 children (6-13)	60	34	41	57	192
9 th Group	57	52	48	55	212
Couples with 1 child (6-13) & 1 child student (14-19)	54	45	41	49	189
Couples with 1 child (6-13) & 1 child student >20	3	7	7	6	23
10 th Group	42	42	29	35	148
Couples with 2 children students (14-19)	14	19	15	10	58
Couples with 1 child student (14-19) & 1 child student >20	17	15	12	19	63
Couples with 2 children students >20	11	8	2	6	27

Households with two children have been aggregated in further five groups, as reported in Table 8. These aggregations are due to practical exigencies to keep the number of possible typologies into a reasonable number and to avoid cells with zero frequencies too.

4. Welfare index-shares

Following Filippucci et al. (2002) we do not restrict the estimation of equivalence scales to models where, according to the original Engel's idea, the dependent variable is the share of total expenditure devoted to food. Besides the usual share of food and non-alcoholic beverages, ${}_1w$, we specify four further shares. The second share, ${}_2w$, is obtained by aggregating to food and non-alcoholic beverages the expenditures which reveal an elasticity less than 1 with respect to total expenditure, when they are regressed one by one in form of share on the logarithm of total expenditure, the logarithm of total family components and year dummies.

These expenditures are:

- Food
- House monthly rent
- Figurative house monthly rent
- Electric power
- Gas
- Telephone
- Oil for Heating
- Central heating plant
- Firewood and coal
- Potable water
- Boys, children and babies clothing
- Children and babies footwear
- Tickets for buses, underground
- Tickets for trains
- Articles for babies
- Driving lessons
- Refectories
- Paper handkerchiefs, napkins, toilet paper
- Exercise books, stationery articles, copies and fax

In the third share, ${}_3w$, in the fourth, ${}_4w$ and in the fifth, ${}_5w$, we add the expenditures which show an elasticity with respect to total expenditure that becomes less than 1, when we introduce the income class dummies, defined in Table 14, into the regression above specified.

Specifically: into ${}_3w$ we add just the expenditures strongly related to the presence of children, so that ${}_3w$ is ${}_2w$ plus the expenditure share (with respect to total expenditure) for the following items:

- School fees
- School charges

- School books
- Lodging and maintenance expenses for living away from one's own family

The fourth share, ${}_4w$ is plus the expenditure share for adult footwear, and the fifth, ${}_5w$, is obtained by adding household's expenditures for transport fuel to the numerator of ${}_4w$.

Table 9. Average shares according to household typologies

	${}_1w$	${}_2w$	${}_3w$	${}_4w$	${}_5w$
Couples without children	0.1494	0.4072	0.4116	0.4232	0.4775
Couples with 1 child <3	0.1495	0.4435	0.4595	0.4664	0.5189
Couples with 1 child 3-5	0.1544	0.4346	0.4539	0.4634	0.5185
Couples with 1 child 6-13	0.1527	0.4249	0.4427	0.4523	0.5025
Couples with 1 child student 14-20	0.1694	0.4326	0.4584	0.4715	0.5236
Couples with 1 child student >20	0.1446	0.3877	0.4389	0.4500	0.4983
6th group households	0.1726	0.4656	0.4857	0.4933	0.5398
7th group households	0.1703	0.4472	0.4697	0.4771	0.5276
8th group households	0.1817	0.4484	0.4726	0.4790	0.5254
9th group households	0.1851	0.4457	0.4770	0.4884	0.5356
10th group households	0.1559	0.4097	0.4489	0.4618	0.5102
Totald dataset	0.1595	0.4267	0.4464	0.4565	0.5075

We are aware that the procedure through which we construct the further four share indicators is rather rough: however we think that it can give some useful indications.

Table 10. Average shares according to household typologies at low income level[4]

Low level income	${}_1w$	${}_2w$	${}_3w$	${}_4w$	${}_5w$
Couples without children	0.1829	0.4682	0.4690	0.4794	0.5378
Couples with 1 child <3	0.1923	0.5104	0.5146	0.5209	0.5828
Couples with 1 child 3-5	0.1821	0.4812	0.4984	0.5078	0.5733
Couples with 1 child 6-13	0.1677	0.4470	0.4642	0.4743	0.5240
Couples with 1 child student 14-20	0.2034	0.4938	0.5182	0.5315	0.5892
Couples with 1 child student >20	0.1795	0.4504	0.4979	0.5065	0.5615
6th group households	0.2027	0.5111	0.5225	0.5282	0.5775
7th group households	0.2004	0.4855	0.4969	0.5045	0.5582
8th group households	0.2055	0.4972	0.5164	0.5224	0.5686
9th group households	0.2256	0.4935	0.5173	0.5268	0.5815
10th group households	0.1853	0.4617	0.4900	0.5030	0.5537

[4] Incomes between the minimum and the 33rd percentile.

Table 11. Average shares according to household typologies at middle income level[5]

Medium level income	$_1w$	$_2w$	$_3w$	$_4w$	$_5w$
Couples without children	0.1338	0.3756	0.3808	0.3928	0.4475
Couples with 1 child <3	0.1293	0.4147	0.4347	0.4426	0.4933
Couples with 1 child 3-5	0.1449	0.4244	0.4408	0.4501	0.5025
Couples with 1 child 6-13	0.1489	0.4178	0.4351	0.4444	0.4971
Couples with 1 child student 14-20	0.1621	0.4145	0.4402	0.4532	0.5053
Couples with 1 child student >20	0.1421	0.3942	0.4431	0.4543	0.5042
6th group households	0.1563	0.4443	0.4665	0.4751	0.5216
7th group households	0.1486	0.4178	0.4465	0.4531	0.5043
8th group households	0.1717	0.4208	0.4458	0.4519	0.4998
9th group households	0.1698	0.4277	0.4566	0.4702	0.5117
10th group households	0.1533	0.4045	0.4475	0.4618	0.5109

Tables 9, 10 and 11 show how the five shares of necessity aggregates vary according household composition and income level. We can observe that all the shares reveal a tendency to decrease with income and to be higher for couples with children. The maximum for food share amounts to 22.6% and is registered in correspondence to low-income level and for couples with two children; the minimum, 9.6%, is attained by high-income level and for couples without children.

Table 12. Average shares according to household typologies at high income level[6]

High level income	$_1w$	$_2w$	$_3w$	$_4w$	$_5w$
Couples without children	0.0958	0.3192	0.3325	0.3461	0.3877
Couples with 1 child <3	0.1040	0.3613	0.3962	0.4008	0.4328
Couples with 1 child 3-5	0.1165	0.3478	0.3851	0.3952	0.4331
Couples with 1 child 6-13	0.1283	0.3947	0.4154	0.4254	0.4670
Couples with 1 child student 14-20	0.1308	0.3704	0.3983	0.4113	0.4555
Couples with 1 child student >20	0.1124	0.3054	0.3657	0.3794	0.4168
6th group households	0.1235	0.3795	0.4226	0.4338	0.4710
7th group households	0.1424	0.4168	0.4549	0.4645	0.5029
8th group households	0.1615	0.4293	0.4622	0.4703	0.5125
9th group households	0.1340	0.3852	0.4408	0.4505	0.4963
10th group households	0.1273	0.3605	0.4036	0.4129	0.4568

$_5w$ exceeds 58% for low-income level typologies with children and registers a minimum of 38.8% with high income level and for couples

[5] Incomes between the 33rd percentile and the 66th percentile.
[6] Incomes between the 66th percentile and the maximum.

without children. Differences between $_2w$ and $_5w$ do not exceed 10-11 percent points, while the step between $_1w$ and $_2w$ reaches somewhere even 30 percent points.

5. Estimation results

We estimate equation (4) applying Box-Cox procedure[7]. After having estimated λ, we transform the dependent variable as it appears in the left hand side of (4) and then we run again ordinary least squares with the option for White's heteroskedasticity-consistent covariance matrix[8].

Vector $\mathbf{d}_h$ contains ten dummies: the first five dummies correspond to the categories listed in the first five rows of Table 7, and the other dummies correspond to the five groups listed in Table 8; this ten dummies are summarized in Table 13.

After many trials we decide to include in vector $\mathbf{z}_h$ the variables, which proved to be significant. These control variables are listed in table 14.

Table 13. Dummy variables included in vector $\mathbf{d}_h$

Couple with 1 child <3
Couple with 1 child 3-5
Couple with 1 child 6-13
Couple with 1 child student 14-20
Couple with 1 child student >20
6th group household
7th group household
8th group household
9th group household
10th group household

We remark that the role of control variables is to take account for possible distortions due to household characteristics: for example, families with a housewife, ceteris paribus, show a share for food higher than where there is not such a presence: this is mainly due to the fact that where there is an housewife more meals are consumed at home than in her absence.

[7] Shazam 9, *User's Reference Manual,* Ch. 12
[8] Shazam 9.

Table 14. Control variables included in vector z_h

House Rented: dummy=1 if the house is rented
Housewife: dummy =1 if in the household there is an housewife
Post-Graduated or Graduated: dummy =1 if the reference person is graduated or post-graduated
Age of the Reference Person: this variable is expressed in years
Low Income: dummy=1 if the household income is between the minimum and the 33rd percentile, 0 otherwise
High Income: dummy=1 if the household income is between the 66th percentile and the maximum, 0 otherwise

Table 15 reports the estimation result obtained through the procedure summarized at the beginning of the paragraph. We stress that the hypothesis λ=1 is strongly rejected by the likelihood-ratio test: moreover, the λ's become closer to one, and decrease in the χ^2 values from ${}_1w$ to ${}_5w$, as the welfare-index share includes a broader set of aggregates. Also the fittings become better as the dimensions of the reference aggregates increase.

It may be interesting to analyse briefly the results obtained for control variables. Ceteris paribus, the demanded food share is higher in the presence of a housewife and increases with the age of the reference person: yet, both variables show coefficients which decrease as the dependent variable aggregates become broader, that is from ${}_1w$ to ${}_5w$.

Having a rented house the share for food raises and decreases the remaining shares: the strongest negative effect is observed for ${}_2w$. Trying to interpret this result we observe that the average monthly amount for rented houses in the sample is 310 Euro, while the average figurative rent is 490 Euro, the latter value is then 59% higher than the former. The figurative rents are estimated directly by house-owners, we think they might report on the form an upper borderline value, or even a value, which may be much greater than the maximum market one. In fact the item for figurative rents is so formulated: "At what price would you rent your house?"; to such a formulation who has no intention to rent his own house, probably answers reporting a quite high price which include both the real and immaterial costs that he should face to leave his residence. Moreover, upward biased values may be due to personal owners' consideration for their own houses.

Table 15. Estimation results of model (4) Box-Cox transform and heteroskedasticity consistent standard errors. Equivalence scales are w.r.t. couple without children.

	1^w		2^w		3^w		4^w		5^w	
	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.
Couple with 1 child <3	1.3053	0.1000	1.3129	0.0499	1.3969	0.0572	1.3567	0.0559	1.3035	0.0519
Couple with 1 child 3-5	1.2961	0.1169	1.2214	0.0440	1.3162	0.0526	1.2938	0.0519	1.2575	0.0475
Couple with 1 child 6-13	1.2883	0.0773	1.2110	0.0343	1.2894	0.0393	1.2736	0.0387	1.2322	0.0362
Couple with 1 child student 14-20	1.5033	0.1243	1.2532	0.0432	1.3749	0.0523	1.3836	0.0518	1.3620	0.0487
Couple with 1 child student >20	1.1268	0.0868	1.0864	0.0368	1.3204	0.0587	1.3227	0.0586	1.3035	0.0542
6^{th} group household	1.6688	0.1421	1.4039	0.0535	1.5374	0.0622	1.4998	0.0607	1.4098	0.0533
7^{th} group household	1.7222	0.1324	1.3833	0.0501	1.5108	0.0600	1.4760	0.0583	1.4165	0.0539
8^{th} group household	1.9103	0.1559	1.3787	0.0476	1.5198	0.0569	1.4811	0.0548	1.4080	0.0502
9^{th} group household	1.9351	0.1496	1.4185	0.0486	1.6009	0.0631	1.5965	0.0629	1.5258	0.0586
10^{th} group household	1.5379	0.1178	1.3372	0.0504	1.5524	0.0694	1.5623	0.0713	1.5331	0.0652
	R^2	λ	R^2	λ	R^2	λ	R^2	λ	R^2	λ
	0.3270	0.4900 (460.18)	0.5962	0.5900 (139.38)	0.5462	0.6700 (78.54)	0.5433	0.6900 (67.14)	0.5627	0.7900 (26.98)
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
(Constant)	0.0973	1.44	1.8471	37.56	1.5694	32.48	1.5655	33.26	1.6175	37.36
Log total expenditure	-0.1957	-25.22	-0.3382	-58.42	-0.2988	-52.49	-0.2940	-52.99	-0.2850	-56.32
Couple with 1 child <3	0.0521	3.54	0.0921	7.21	0.0999	8.28	0.0897	7.49	0.0755	6.71
Couple with 1 child 3-5	0.0507	2.93	0.0676	5.58	0.0821	6.95	0.0757	6.48	0.0653	6.11
Couple with 1 child 6-13	0.0496	4.29	0.0647	6.79	0.0759	8.42	0.0711	8.02	0.0595	7.13
Couple with 1 child student 14-20	0.0798	4.93	0.0763	6.55	0.0951	8.46	0.0954	8.76	0.0880	8.71
Couple with 1 child student >20	0.0234	1.54	0.0280	2.44	0.0830	6.27	0.0822	6.32	0.0755	6.38
6^{th} group household	0.1002	6.14	0.1147	9.03	0.1285	10.91	0.1192	10.25	0.0979	9.28
7^{to} group household	0.1064	7.41	0.1097	9.06	0.1233	10.63	0.1145	10.06	0.0992	9.28
8^{th} group household	0.1266	8.37	0.1086	9.32	0.1251	11.33	0.1155	10.73	0.0975	9.66
9^{th} group household	0.1292	8.72	0.1183	10.18	0.1406	12.01	0.1375	11.98	0.1204	11.10
10^{th} group household	0.0842	5.62	0.0983	7.68	0.1314	9.89	0.1312	9.81	0.1218	10.05
1997	-0.0370	-3.93	-0.0331	-4.43	-0.0256	-3.53	-0.0276	-3.86	-0.0339	-5.11
1998	-0.0390	-3.95	-0.0287	-3.63	-0.0241	-3.13	-0.0243	-3.21	-0.0219	-3.15
1999	-0.0230	-2.31	-0.0241	-3.06	-0.0158	-2.07	-0.0171	-2.27	-0.0176	-2.51
Households living in a rented house	0.0329	3.23	-0.0515	-6.67	-0.0409	-5.60	-0.0422	-5.89	-0.0377	-5.66
Housewife	0.0302	3.43	0.0301	4.30	0.0224	3.22	0.0222	3.25	0.0208	3.30
Post-graduated / graduated	-0.0339	-3.34	0.0260	3.23	0.0337	4.09	0.0340	4.23	0.0239	3.29
Age of the reference person	0.0037	7.21	0.0023	5.64	0.0022	5.73	0.0020	5.19	0.0014	3.78
Low income households	0.0307	3.34	-0.0062	-0.86	-0.0068	-0.98	-0.0081	-1.20	-0.0081	-1.30
High income households	-0.0073	-0.75	0.0307	3.68	0.0362	4.22	0.0368	4.36	0.0312	4.08

In parentheses log-likelihood ratio χ^2 statistic testing H_0: λ =1 against H_0: $\lambda \neq 1$

When figurative house rents are upward biased, if their value is added to total expenditure and is not added to the share numerator, as it happens for ${}_1w$, the shares become smaller than they should. In such a situation, as the effect of figurative rent inflation has less effect on the logarithm of total expenditure than on the denominator of ${}_1w$, ceteris paribus, the logarithm of total expenditure gives a fitted value for the dependent variable higher than the actually registered one. To compensate this phenomenon, the constant term is probably pushed down for house owners, which, conversely, implies positive dummy coefficient for households who live in rented houses.

On the contrary, when an upward biased value is added both to total expenditure and to the numerator of the share, as it happens for the remaining four shares, the shares tends to be greater than they should, and the logarithm of total expenditure tends to generate fitted values for the dependent variable lower than those which are actually registered: so, to compensate positive fitting errors, the constant term for house-owners is pushed up. In turns this implies negative dummies for house-renters. The effect is expected to be more evident when the overestimated figurative rents are relatively greater with respect to the value of the correct amount at the numerator: actually the dummy coefficient for house-renters show its maximum absolute value for ${}_2w$ and decreases as long as more items are added to the numerator of the share, as it happens passing from ${}_2w$ to ${}_3w$, ${}_4w$ and ${}_5w$.

In what it concerns income dummies, we notice that the 48,5% of house renters are in the low income class, the 41,2% in the middle class and the 10.3% in the high; conversely the 35.3% of house-owners are in the low income class, the 47.5% in the middle and the 17.2% is in the high income class. Then in the considered sample house-owners are shifted to higher income classes than house-renters: so these income dummies have a role similar and complementary to that of the house renter dummy: this may explain why the low income dummy is positive for ${}_1w$ while the high income one is negative. Conversely, when the dependent variable is ${}_2w$, ${}_3w$, ${}_4w$ and ${}_5w$, it may be the over-evaluation of figurative rents which makes positive the coefficient of the high income level dummy and negative that of the low income dummy.

Of course we cannot explain the whole behaviour of the income dummies with over evaluations of figurative rents: these dummies can possibly be proxies for other unspecified household characteristics, habits and variables.

The estimated scales coefficients are reported in the upper part of Table 15: these coefficients are calculated w.r.t. the couple without children, which is set to 1. All the estimated scales are highly significant. Figure 1 shows how the scale coefficients vary in relation to the ten family typologies with children and according to the five share indicators adopted. The graph reports both point estimates for the scales and interval estimates obtained by adding and subtracting to each point estimates its standard error.

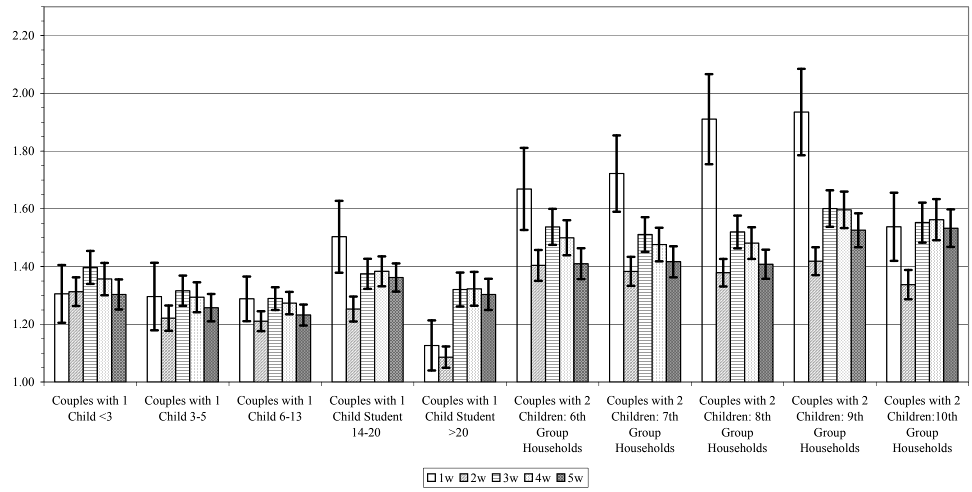

Fig. 1. Model (4): scale coefficients for the ten households typologies. The vertical line at the top of each stick represents the interval: "scale coefficient estimate ± its standard error"

Looking at Table 15 and Figure 1 we can observe that scale coefficients obtained by ${}_1w$ have larger standard errors than those related to the remaining shares. Food gives scales similar to those given by ${}_3w$, ${}_4w$ and ${}_5w$ for household typologies represented in group 1,2 and 3; that is for families with one child less than 13, and for group 10. The food scale is higher than the others when there is a student between 14 and 20; it is definitely much higher than the remaining scales in correspondence of family groups 6, 7, 8 and 9. On the contrary, when there is one student older than twenty, ${}_1w$ gives a very low scale if compared to those of ${}_3w$, ${}_4w$ and ${}_5w$: this is probably due to the fact that students older than twenty eat at home rather seldom with respect to younger children.

In any case, when there is a student older than 20, we would have expected scale coefficients notably higher than those obtained for groups 1, 2, 3 and 4.

We observe that, if we exclude families with a small child, 2w is the share, which gives the lowest scales.

Costs for one child seem to be decreasing with child's age until when he/she is thirteen: the scale coefficients, range between 1.30 to 1.40 for a child less than three years old, according to the share adopted as indicator; they range between 1.22 to 1.30 when children are between 3 and five, between 1.21 to 1.29 when they become between 6 and 13 years old.

Costs increase markedly when children becomes from fourteen to twenty years old: for this group scale coefficients range between 1.25 and 1.50. As we have already observed, costs decrease unexpectedly when sons or daughters are more than twenty: this reduction is probably due to the fact that when sons or daughters are more then twenty they buy by themselves many things that they need, even if by money received by parents. To get more realistic costs for this group ISTAT's survey should get information also for money directly transferred from parents to children.

Economies of scale can be directly observed only comparing group 3 and group 8, that is families with one child between 6 and 13 and families with two children both between 6 and 13 years old. If we rely on food, the presence of a further child in this age class should even induce scale diseconomies. According to the other shares costs for a further child increase less proportionally: however these scale economies are rather slight, as they are less than 6 per cent points.

If we want to get a synthetic measure of scale economies, which households achieve with a second child, we can compare the differences between the average scale coefficients for households with one child and the corresponding averages obtained for households with two children. The averages for each aggregate are reported in Table 16. From it we observe that in what it concerns food the presence of a further child would increase costs more then proportionately with respect to family with one child. With respect to the other shares, costs increase less than proportionately. These are the average scale economies: 5 per cent points for ${}_2w$, 12 per cent points for ${}_4w$ and ${}_5w$ and 14 per cent points for ${}_3w$; we observe that if we do not consider ${}_1w$, ${}_3w$ gives the highest scale coefficients.

Table 16. Average scale coefficients for households with one child and with two children

	${}_1w$	${}_2w$	${}_3w$	${}_4w$	${}_5w$
Couple with 1 child	1.31	1.22	1.34	1.32	1.29
Couple with 2 children	1.76	1.39	1.54	1.52	1.46
Couple with 2 children: no scale effects	1.62	1.44	1.68	1.64	1.58

6. Considerations on the specification effects

In this section we compare the results obtained through two different specifications of Engel equation. These two specifications are given at equation (6) and equation (7)

$$ {}_g w_h = \alpha + \beta \log y_h + \boldsymbol{\theta}'\mathbf{d}_h + \tau_1 t_{97} + \tau_2 t_{98} + \tau_3 t_{99} + \varepsilon_h , \qquad (6) $$

$$ {}_g w_h = \alpha + \beta \log y_h + \boldsymbol{\theta}'\mathbf{d}_h + \boldsymbol{\psi}'\mathbf{z}_h + \tau_1 t_{97} + \tau_2 t_{98} + \tau_3 t_{99} + \varepsilon_h . \qquad (7) $$

Model (7) differs from model (4) just in the l.h.s. of the equation: (7) can be obtained from (4) if we restrict λ to 1. (6) is obtained in turn from (7) if we omit the vector of control variables **z**.

The estimation results for the two models are reported, respectively, in Table 17 and Table 18. Scale coefficients for model (6), (7) and (4) can be easier compared looking at figures 2, 3, 4, 5, 6.

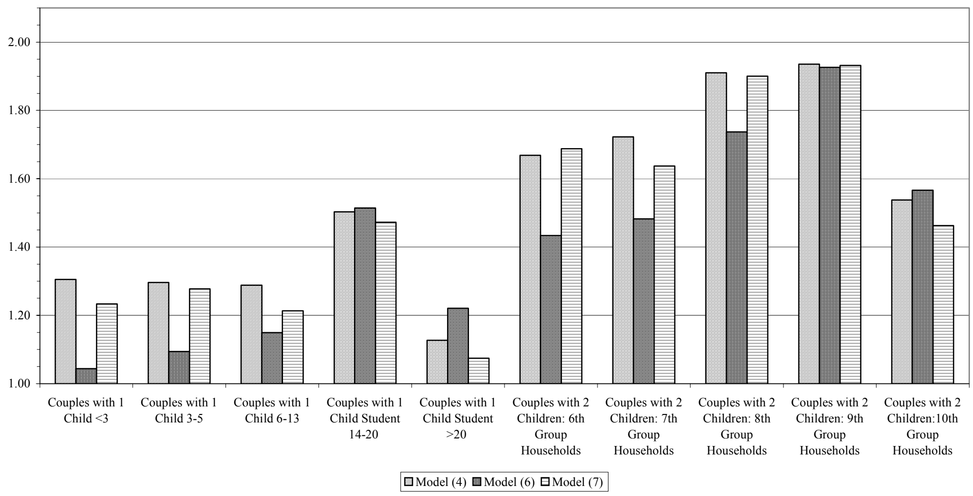

Fig. 2. Share ${}_1w$

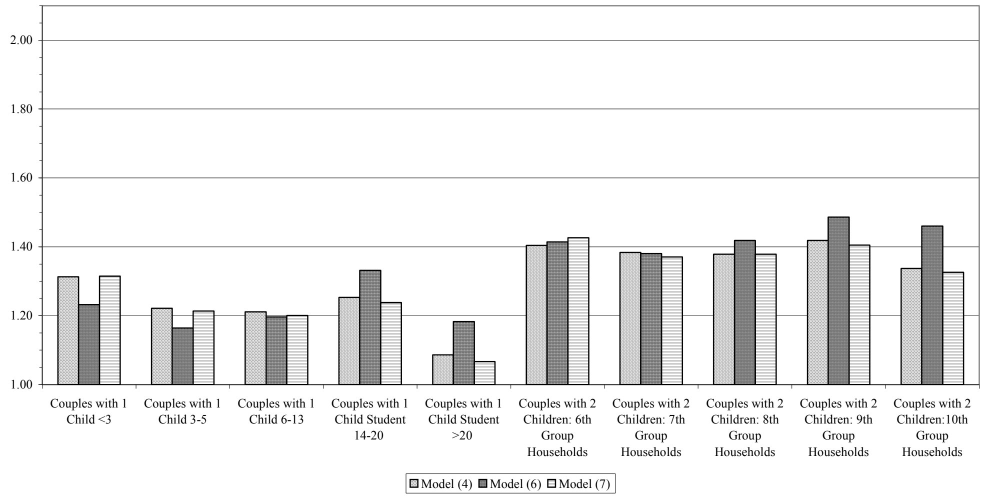

Fig. 3. Share $_2w$

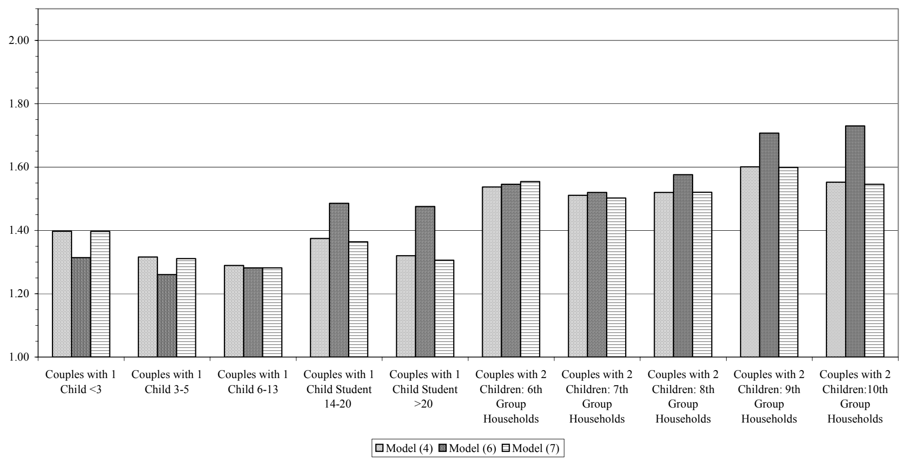

Fig. 4. Share $_3w$

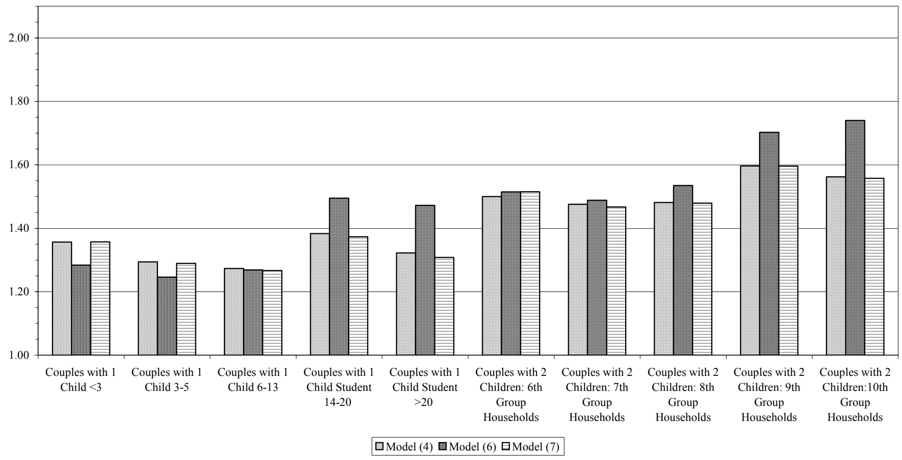

Fig. 5. Share $_4w$

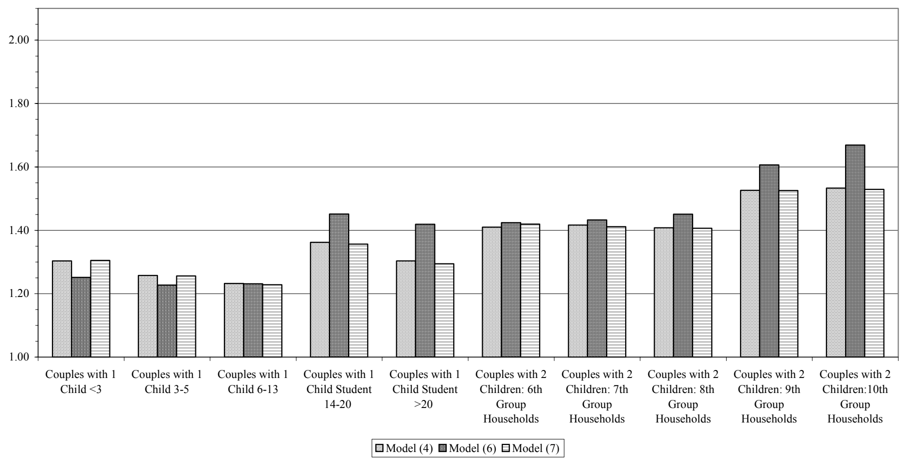

Fig. 6. Share $_5w$

Table 17. Estimation results of model (6)

Model 1	$_1w$		$_2w$		$_3w$		$_4w$		$_5w$	
	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.
Couple with 1 child <3	1.0436	0.0682	1.2319	0.0503	1.3145	0.0578	1.2837	0.0568	1.2518	0.0521
Couple with 1 child 3-5	1.0941	0.0872	1.1640	0.0457	1.2608	0.0548	1.2461	0.0543	1.2275	0.0494
Couple with 1 child 6-13	1.1492	0.0629	1.1967	0.0382	1.2816	0.0440	1.2689	0.0437	1.2314	0.0399
Couple with 1 child student 14-20	1.5145	0.1140	1.3316	0.0511	1.4860	0.0627	1.4951	0.0623	1.4513	0.0557
Couple with 1 child student >20	1.2207	0.0832	1.1828	0.0415	1.4759	0.0699	1.4723	0.0697	1.4189	0.0609
6th group household	1.4335	0.1183	1.4142	0.0585	1.5461	0.0677	1.5147	0.0668	1.4241	0.0570
7th group household	1.4820	0.1074	1.3803	0.0557	1.5201	0.0667	1.4888	0.0651	1.4325	0.0587
8th group household	1.7368	0.1344	1.4184	0.0556	1.5758	0.0673	1.5351	0.0649	1.4506	0.0573
9th group household	1.9261	0.1502	1.4861	0.0554	1.7072	0.0754	1.7031	0.0751	1.6064	0.0671
10th group household	1.5665	0.1058	1.4603	0.0572	1.7301	0.0820	1.7401	0.0843	1.6687	0.0736
	R^2		R^2		R^2		R^2		R^2	
	0.2725		0.5535		0.5064		0.5046		0.5365	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
(Constant)	0.7952	35.86	2.1151	71.67	2.0376	64.84	2.0593	65.63	2.2730	71.89
Log total expenditure	-0.0818	-29.90	-0.2168	-58.97	-0.2066	-52.63	-0.2077	-53.00	-0.2278	-57.55
Couple with 1 child <3	0.0035	0.65	0.0452	5.13	0.0565	6.28	0.0519	5.69	0.0512	5.44
Couple with 1 child 3-5	0.0074	1.13	0.0329	3.88	0.0479	5.37	0.0457	5.08	0.0467	5.12
Couple with 1 child 6-13	0.0114	2.56	0.0389	5.64	0.0512	7.27	0.0495	6.94	0.0474	6.43
Couple with 1 child student 14-20	0.0339	5.52	0.0621	7.47	0.0818	9.44	0.0835	9.70	0.0849	9.73
Couple with 1 child student >20	0.0163	2.91	0.0364	4.76	0.0804	8.22	0.0804	8.18	0.0797	8.15
6th group household	0.0294	4.35	0.0751	8.47	0.0900	10.17	0.0862	9.60	0.0805	8.98
7ho group household	0.0322	5.52	0.0699	8.05	0.0865	9.71	0.0827	9.24	0.0819	8.87
8th group household	0.0451	7.29	0.0758	8.94	0.0939	10.78	0.0890	10.25	0.0847	9.49
9th group household	0.0536	8.60	0.0859	10.66	0.1105	12.29	0.1106	12.27	0.1080	11.51
10th group household	0.0367	6.68	0.0821	9.62	0.1132	11.56	0.1151	11.42	0.1166	11.55
1997	-0.0105	-2.79	-0.0248	-4.64	-0.0219	-3.91	-0.0242	-4.28	-0.0318	-5.48
1998	-0.0113	-2.82	-0.0184	-3.22	-0.0174	-2.91	-0.0184	-3.05	-0.0188	-3.07
1999	-0.0068	-1.67	-0.0155	-2.72	-0.0107	-1.78	-0.0121	-2.01	-0.0144	-2.35

Table 18. Estimation results of model (7)

Model 2	${}_1w$		${}_2w$		${}_3w$		${}_4w$		${}_5w$	
	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.	Scale	Scale S.E.
Couple with 1 child <3	1.2332	0.0957	1.3148	0.0527	1.3968	0.0591	1.3570	0.0575	1.3053	0.0529
Couple with 1 child 3-5	1.2777	0.1182	1.2134	0.0458	1.3109	0.0540	1.2891	0.0530	1.2560	0.0484
Couple with 1 child 6-13	1.2133	0.0751	1.2007	0.0356	1.2816	0.0405	1.2669	0.0399	1.2285	0.0370
Couple with 1 child student 14-20	1.4726	0.1288	1.2383	0.0449	1.3640	0.0539	1.3737	0.0534	1.3568	0.0496
Couple with 1 child student >20	1.0746	0.0901	1.0667	0.0381	1.3063	0.0611	1.3085	0.0608	1.2944	0.0556
6th group household	1.6878	0.1619	1.4262	0.0580	1.5547	0.0657	1.5156	0.0638	1.4193	0.0546
7th group household	1.6372	0.1362	1.3706	0.0524	1.5023	0.0619	1.4673	0.5977	1.4116	0.0548
8th group household	1.9004	0.1715	1.3784	0.0502	1.5205	0.0595	1.4795	0.5692	1.4061	0.0512
9th group household	1.9318	0.1697	1.4051	0.0504	1.5993	0.0659	1.5961	0.0655	1.5255	0.0599
10th group household	1.4628	0.1200	1.3260	0.0513	1.5460	0.0713	1.5578	0.0735	1.5290	0.0665
	R^2		R^2		R^2		R^2		R^2	
	0.3084		0.5761		0.5300		0.5283		0.5538	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
(Constant)	0.6626	24.88	2.1158	61.93	2.0566	57.19	2.0923	58.48	2.3103	63.52
Log Total Expenditure	-0.0710	-23.85	-0.2255	-56.29	-0.2182	-51.81	-0.2203	-52.56	-0.2397	-56.60
Couple with 1 child <3	0.0149	2.75	0.0617	6.87	0.0729	8.01	0.0673	7.28	0.0639	6.63
Couple with 1 child 3-5	0.0174	2.69	0.0436	5.15	0.0591	6.64	0.0560	6.24	0.0546	5.96
Couple with 1 child 6-13	0.0137	3.17	0.0412	6.19	0.0541	7.93	0.0521	7.57	0.0493	6.86
Couple with 1 child student 14-20	0.0275	4.46	0.0482	5.92	0.0677	7.95	0.0700	8.26	0.0731	8.42
Couple with 1 child student >20	0.0051	0.86	0.0146	1.81	0.0583	5.73	0.0592	5.80	0.0619	6.03
6th group household	0.0372	5.53	0.0800	8.84	0.0963	10.70	0.0916	10.09	0.0839	9.29
7th group household	0.0350	6.18	0.0711	8.35	0.0888	10.11	0.0845	9.60	0.0826	9.02
8th group household	0.0456	7.42	0.0724	8.83	0.0914	10.84	0.0863	10.28	0.0817	9.41
9th group household	0.0468	7.74	0.0767	9.49	0.1024	11.50	0.1030	11.53	0.1012	10.87
10th group household	0.0270	4.65	0.0636	7.28	0.0950	9.51	0.0977	9.44	0.1018	9.78
1997	-0.0129	-3.52	-0.0228	-4.36	-0.0195	-3.53	-0.0216	-3.88	-0.0293	-5.11
1998	-0.0128	-3.28	-0.0180	-3.25	-0.0167	-2.86	-0.0176	-2.98	-0.0180	-3.00
1999	-0.0075	-1.87	-0.0159	-2.88	-0.0114	-1.96	-0.0128	-2.18	-0.0149	-2.47
Households living in a rented house	0.0156	3.83	-0.0344	-6.36	-0.0301	-5.43	-0.0319	-5.74	-0.0318	-5.55
Housewife	0.0124	3.51	0.0220	4.40	0.0181	3.39	0.0182	3.38	0.0185	3.39
Post-graduated /graduated	-0.0114	-3.24	0.0169	3.06	0.0245	3.96	0.0252	4.09	0.0196	3.18
Age of the reference person	0.0013	6.24	0.0015	5.32	0.0016	5.46	0.0015	4.97	0.0011	3.69
Low income households	-0.0145	4.06	-0.0005	-0.11	-0.0025	-0.48	-0.0038	-0.73	0.0052	-0.96
High income households	-0.0173	-0.83	0.0209	3.72	0.0262	4.13	0.0273	4.26	0.0311	4.01

We observe first of all that model (4) and (7) give rather similar results for shares ${}_2w$, ${}_3w$, ${}_4w$ and ${}_5w$. More evident differences can be observed for food share, even not everywhere.

In fact, from table 15, we remember that the estimates for λ becomes less distant from 1 as the share aggregates become broader: λ is 0.49 for food, 0.59 for ${}_2w$, 0.67 for ${}_3w$, 0.69 for ${}_4w$ and 0.79 for ${}_5w$.

We can conclude that the omission of significant control variables has undoubtedly stronger effects than an incorrect transformation for the dependent variable: also in this case the most evident differences are observed for food.

We notice that generally model (6) underestimates scales for younger groups of children and overestimates scales for families with older children: this bias depends in great part on the omission of the reference person age, which is likely correlated with children's age.

7. Conclusions

The results obtained in this research show that in estimating equivalence scales through Engel curves it is very important to specify in the model suitable control variable to take account for heterogeneity among families.

Though the food share can still be a valid indicator to compare family welfare and then to estimate equivalence scales, in some situations it results to be misleading: this the case for families with a student more than 20 years old.

The experiment of adopting broader good aggregates as indicators gives interesting results and more likely estimates in correspondence of some family typologies: still more investigation is necessary to confirm the results here obtained. For example we must observe that even if, for students more than 20, the scales obtained by broader aggregates seems to be more realistic than that we can estimate referring just to food, it is not convincing that they present lower value than in the case of families with younger children.

Acknowledgements

We desire to thank dr Zofia Rusnak and mgr Marek Kosny for their precious suggestions. Of course all errors and mistakes that should be present are on our own responsibility.

References

[1] Carbonaro G. (1985), *Nota sulla Scala di Equivalenza*, Presidenza del Consiglio dei Ministri, Commissione di Indagine sulla Povertà, *La Povertà in Italia - Studi di Base*, Roma.
[2] Deaton A. and J. Muellbauer (1980), *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge.
[3] Engel E. (1895), *Die Lebenskosten Belgischer Arbeiter-Familien Früher und Jetzt* – „Bulletin de l'Institut International de Statistique" vol. 9.
[4] Filippucci C. and I. Drudi (2002), *Il costo dei figli e dei genitori anziani*, „Famiglie: mutamenti e politiche sociali", Edited by Osservatorio nazionale sulle famiglie e le politiche locali di sostegno alle responsabilità familiari, Chapter 10, Il Mulino, Bologna.
[5] Perali F. (2002), *Some curiosities about the Engel method to estimate equivalence scales,* "*Economic Bulletin*", IV.
[6] Vernizzi A., Rusnak Z. (2002), *Equivalence scales for biological type of households empirical research*, „Badania Operacyjne i Decyzje",.
[7] Vernizzi A., Rusnak Z., Siletti E. (2003), *A contribution to the estimation of equivalence scales in Poland*, „Quality of Life Research", red. W. Ostasiewicz, Chapter 9, Yang's Scientific Press, Tucson (USA).
[8] Whistler D., K.J. White, S.D. Wong and D. Bates (2001), *Shazam: The Econometrics Computer Program*, version 9, Northwest Econometrics, Ltd, Vancouver.

Estymacja kosztów posiadania dzieci za pomocą krzywych Engla

Streszczenie

W artykule dokonano estymacji skal ekwiwalentności w odniesieniu do kosztów posiadania dzieci w jednym z najlepiej rozwiniętych regionów Włoch: Lombardii. Wykorzystane dane pochodzą z ankiet dotyczących wydatków konsumpcyjnych pochodzących z głównego włoskiego urzędu statystycznego (ISTAT). W artykule wykorzystano dane z lat 1997, 1998, 1999 i 2000. Należy podkreślić, że ankiety ISTAT nie mają charakteru panelowego.