

Consumption Behaviour across Regions: a Multilevel Approach

I Comportamenti di Consumo a Livello Regionale: un Approccio Multilevel

Filippa Bono, Miranda Cuffaro, Rosa Giaimo
Dipartimento di Contabilità Nazionale ed Analisi dei Processi Sociali,
Viale delle Scienze, 90128 Palermo
bono@economia.unipa.it; cuffaro@unipa.it; giaimo@unipa.it

Riassunto: L'analisi del comportamento di consumo si è spinta sempre più verso l'utilizzo di micro-dati, relativi alle famiglie, e ciò per tener conto della molteplicità di fattori che guidano le scelte familiari e che vanno a comporre, in ultima analisi, la formazione di decisioni aggregate. In realtà, però, le analisi micro trascurano la composizione delle famiglie in termini di differenti gruppi omogenei di "agenti", in relazione al livello di reddito, alla composizione familiare, alla differente residenza geografica e così via. Per quanto detto, se consideriamo, per esempio, le famiglie nell'ambito del loro contesto di riferimento (le regioni) riproduciamo una struttura gerarchica, che può essere analizzata con un modello multilevel. Vengono stimate, pertanto, delle funzioni "engeliane" di consumo con variabili esplicative relative alle caratteristiche familiari.

Keywords: Consumption patterns, Random parameters model

1. Introduction

Consumer behaviour, inequality and welfare measurements have dominated the empirical literature since the early studies of Engel and Working. The traditional aggregate approach, which treats all households as a single behavioural unit, can be misleading in so far as it ignores the differential behaviour of its constituent micro units. On the other hand, micro-level analysis ignores the natural tendency of data to cluster in homogeneous groups. For example, households living in the same location (or region) are exposed to the same historical, social, economic and environmental factors, or households belonging to the same income class or, those of the same size, behave in a homogeneous way, with similar consumption patterns (Deaton, 1992; Ferrari, 2003). If we consider the statistical units (the households, in this case) within their own context (the regions) a hierarchical structure is reproduced which can be used as a statistical tool applied to hierarchical data. In this paper, a multilevel approach is proposed in order to analyse the consumption behaviour of Italian households in 2000. The households, from Istat-BF (Family Budget), constitute the first level of analysis and are grouped according to their living context: region of residence (the second level). The paper is organised as follows: section 2 briefly describes the basic multilevel model; section 3 presents data and variables and it discusses the empirical findings and section 4 contains some initial conclusions.

2. Multilevel Models

Suppose we have N statistical units in J groups. In order to simultaneously describe the relationship among several households we can write

$$y_{ij} = \alpha_j + \beta_j x_{ij} + e_{ij} \quad (1)$$

where j refers to the level 2 unit and i to level 1 unit. If we let α_j , β_j , become random variables, we obtain a multilevel model. We replace α_j with β_{0j} and β_j with β_{1j} and we assume that $\beta_{0j} = \beta_0 + u_{0j}$ and $\beta_{1j} = \beta_1 + u_{1j}$ where u_{0j} , and u_{1j} are random variables with zero mean and constant variances σ_{u0}^2 , σ_{u1}^2 and covariance σ_{u01} .

We can write (1) in the form

$$\begin{aligned} y_{ij} &= \beta_0 + \beta_1 x_{ij} + (u_{0j} + u_{1j} x_{ij} + e_{0ij}) \\ \text{var}(e_{0ij}) &= \sigma_{e0}^2 \end{aligned} \quad (2)$$

We have expressed the response variable y_{ij} as the sum of a fixed and a random part inside the brackets. The random variables are referred to as 'residuals' and, in the case of a single level model, the level 1 residual e_{0ij} becomes the usual linear model residual term. Equation (2) requires the estimation of two fixed coefficients, β_0, β_1 , and four other parameters, $\sigma_{u0}^2, \sigma_{u1}^2, \sigma_{u01}$ and σ_{e0}^2 . We refer to such variances and covariances as *random parameters*. When the intercept alone is a random parameter we have a particular multilevel model which is named variance components model because the variance of the response about the fixed predictor is the sum of level 1 and level 2 variances (Goldstein, 2003; Hox, 1995).

For the consumption data we test the statistical significance of a random parameters model.

3. Data and results

In order to perform the analysis, we considered the following shares of consumption expenditure: the food expenditure (Q1), the “working” expenditure (Q2), the luxury one (Q3). The former is that which has been conventionally employed in well-being and poverty studies; Q2 includes all the expenditures for those primary needs, excluding food (eg. housing, energy, transports and communications, clothing and health); the latter only includes education, leisure, entertainment and other goods. We commenced our analysis by providing a brief description of the shape of these shares across 16 income classes (proxied by total expenses). Q1 behaves like a semi-logarithmic function, decreasing sharply from the first class to the fourth but slightly decreasing from that one to the fifteenth class; Q2 displays non linear behaviour and it can be approximated by a sigmoid function with an horizontal asymptote starting from the fourth class; on the other hand Q3 exhibits a linear trend, increasing notably between the fourth and the fifteenth class. In order to evaluate the different propensity to

consume, three different engelien consumption function were estimated according to place of residence. The hierarchical structure of data is represented by 23728 households, nested within the nineteen Istat-BF regions (clusters)⁽¹⁾. Focusing on region's variability of shares, we note that median values of Q1 in the nineteen clusters vary from 0.17 to 0.28; those of Q2 from 0.64 to 0.73, meanwhile, those of Q3 from 0.03 to 0.08. The (standardized) shares (i.e. dependent variables) are respectively a semilog, a sigmoid and a linear function of total consumption (as a proxy of income) and of some dummies representing the households characteristics: gender, marital status, job status, education level and the household size centered about its mean of 2,78.

In order to estimate the parameters of models we employed iterative generalized least squares (IGLS) provided by MLwiN software (Rasbash et al., 2000). The main results are summarized below in Table 1.

Table 1: Random coefficient model for Q1, Q2, Q3

	Q1	Lg(Q2)	Q3
<i>Parameter</i>	β (Standard error)	β (Standard error)	β (Standard error)
Constant	-0,239 (0,049)	0,002 (0,043)	0.379 (0,033)
total consumption (²)	-0,496 (0,019)	-0,112 (0,020)	0.238 (0,016)
Gender (male)			
Female	0,021 (0,019)	0,161 (0,023)	-0.227 (0,022)
Marital Status (couple)			
Single	-0,195 (0,020)	0,011 (0,024)	0,113 (0,023)
Job Status (dependent)			
Independent	-0,016 (0,017)	0,002 (0,021)	0,018 (0,020)
Other	0,157 (0,014)	-0,029 (0,017)	-0,180 (0,016)
Educational level (degree)			
High school	0,114 (0,021)	0,027 (0,026)	-0,187 (0,025)
Elementary	0,290 (0,021)	-0,029 (0,025)	-0,379 (0,024)
Household size	0,144 (0,006)	-0,060 (0,008)	-0,031 (0,007)
Level 2 σ_{u0}^2 (intercept)	0,038 (0,013)	0,023 (0,008)	0,009 (0,003)
σ_{u01} (covariance)	-0,012 (0,005)	0,006 (0,003)	-0,002 (0,002)*
σ_{u1}^2 (total consumption)	0,006 (0,002)	0,006 (0,002)	0,003 (0,002)*
Level 1 σ_e^2	0,623 (0,006)	0,938 (0,009)	0,893 (0,009)
ICC	0,062	0,024	0,010

* the χ^2 test for the overall comparison of this model with respect to the model without two variance component is highly significant (D=37, g.d.l.=2, p=0,001).

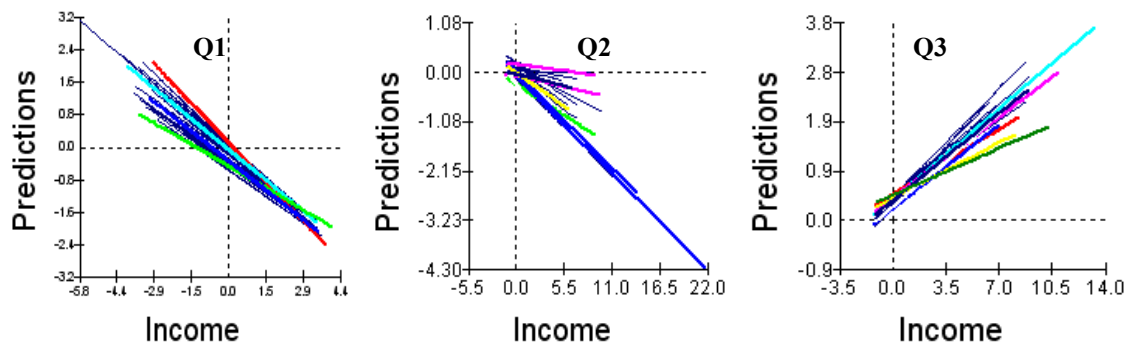
The semilog, sigmoid and linear functions, which were chosen according to the share trend across income classes, achieve the best fit for Q1, Q2, Q3 respectively and the parameters generally possess the right signs. Furthermore, the evidence supports the general strong effect of the household's characteristic variables especially regarding Q3, with a negative and statistically significant impact of gender, job status, education level and size on that share. All the functions show significant, even if very low, values for second level variance estimates. A close examination of Figure 1 reveals a higher

⁽¹⁾ The dimension of clusters vary from 516 to 2430 households.

⁽²⁾ Measured respectively by: log of standardized variable for Q1, inverse standardized variable for LgQ2, standardized variable for Q3.

regional variability in the propensity to consume Q1 and Q2 at lower incomes (below the mean) and a propensity to consume Q3 at higher incomes respectively. We can, therefore, reasonably suppose that the difference in the regional consumption pattern of primary goods (food and “working” share) is determined by different price levels and eating habits. These factors have a strong influence on the consumption patterns of low-income households. Conversely, the difference relating to Q3 are due more to social and demographic characteristics than to differences in the cost of living. In either case, our results stress the major role of income as compared with context and they can be considered as a valid starting point for further investigation.

Figure 1: Prediction lines at level 2 for Q1, Q2 and Q3- casual β_0 and β_1



4. Conclusion

This preliminary analysis, which was carried out on regional consumption patterns obtained by a multilevel approach, contains some interesting results that could make this field worth exploring. On the one hand, the multilevel model revealed itself to be a useful tool to explore such type of data and complex relationships; on the other hand, the results indicate that the income distribution within regions has a strong effect on consumption patterns and it could, therefore, be considered in terms of a multilevel cross-classification model, by which households may be classified according to two dimensions: regional and income class.

References

- Deaton A. (1992) *Understanding Consumption*, Clarendon Press, Oxford.
- Ferrari G. (2003) Using equivalence scales as spatial deflators: evidence in inter-households welfare regional comparisons from italian HBS micro-data, in *Household Behaviour, Equivalence Scales, Welfare and Poverty*, C.Dagum and G.Ferrari eds. Physica-Verlag.
- Goldstein H. (2003) *Multilevel Statistical Models*, Oxford University Press, Inc., NY.
- Hox J. (1995) *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties.
- Rasbash J., Browne W., Goldstein H., et al. (2000) *A User's Guide to MLwiN*, 2nd edition. London, Institute of Education.