UNIVERSITAT POMPEU FABRA
Facultat de Ciències Econòmiques i Empresarials

Departament d'Economia

Demand Analysis of Non-durable Goods Relaxing Separability Assumptions

Jordi Puig Gabau

1998

UNIVERSITAT POMPEU FABRA
Facultat de Ciències Econòmiques i Empresarials

Departament d'Economia

Demand Analysis of Non-durable Goods. Relaxing Separability Assumptions.

Jordi Puig Gabau

1998

El present treball per optar al grau de doctor ha estat realitzat sota la direcció d'en José María Labeaga Azcona, Professor Titular de la UNED.

Als meus pares.

CONTENTS

AGRAÏMENTS

Mitjançant aquestes línies, voldria donar les gràcies a tota aquella gent que ha estat al meu costat al llarg d'aquest temps i que, d'una o altra forma, m'han ajudat a fer aquesta tesi doctoral.

En primer lloc, vull expressar la meva més sincera gratitud al meu supervisor, en José María Labeaga no només pel seu ajut i guia intellectual, sinó també pels seus ànims i suport moral. També a l'Angel López, el meu tutor, per tots els comentaris i suggerències sobre aquest treball. A en Jaume García per introduir-me en el camp de la recerca en Economia Aplicada. I al Departament d'Economia de la Universitat Pompeu Fabra per posar l'entorn i la motivació per començar a fer investigació.

Part d'aquest treball s'ha presentat en vàries conferències i seminaris. Agraeixo als assistents els seus comentaris, i molt especialment, a Martin Browning i Jean-Marc Robin les seves encertades suggerències.

També vull recordar tota aquella gent molt propera que sempre m'ha recolzat. Vull mencionar especialment a en Carles, en Joan, la Marta, l'Eulàlia, en Màrius, la Mireia, la Reyes i a tota la gent de l'anomenada generació zero. Tots vosaltres hi teniu miquetes en aquest treball.

Vull agrair també a tota la meva família, i molt especialment als meus pares, el que hagin estat sempre aquí. Són ells qui m'han donat un esperit de lluita i de treball que ha estat el meu principal suport.

Finalment, vull agrair l'ajut financer proveït pel Ministerio de Educación y Ciencia (DGICYT PB95-0980).

LIST OF TABLES AND FIGURES

CHAPTER 1 - SUMMARY AND CONCLUSIONS

1.1 - THEORETICAL BACKGROUND

Consumers choose upon income and leisure and also take decisions about the allocation of the former between consumption and savings. Moreover, the process includes the assignment of expenditure between different consumption goods. Thereby, the whole allocation problem implies interaction among decisions taken in different issues. Translating this problem to a temporal perspective, we may also consider the interaction of today's choices with decisions located in the following future. The specification and analysis of all the implied relationships is certainly too broad. Usually, the aim and interest of the researcher implies the introduction of restrictions on the way decisions within a period or among different periods are taken.

The specification of consumer preferences provides the framework for undertaking the simplification of such a large allocation problem. Besides, separability assumptions on preferences may characterize, in simpler ways, how do consumers proceed on their decisions and choices on income, leisure, consumption or savings. There are different types of separability which may define the decision processes (Pudney, 1981).

We consider the different types of separable structures of consumer preferences by defining first a vector of consumption goods $Q=(Q_1, \ldots, Q_N)$ where each argument, Q_G, is an aggregate of different commodity items, and a utility function $V(\cdot)$ representing strictly convex preferences.

a) Additive or strong separability.

Consumer preferences are additive or strongly separable among goods if consumption on a particular good is not influenced by consumption on items included in other goods. In this case, the utility function $V(\cdot)$ is constructed as a sum of subutility functions for each good. Hence, utility must be formulated with the following expression:

[1.1]

$$V(Q_1,\ldots,Q_N)=V_1(Q_1) + ,\ldots, + V_N(Q_N).$$

The implications from this formulation are that the marginal utilities on the different groups are independent which means that $\partial V_G(Q_G) / \partial q_j = 0$ if $q_j \notin Q_G$.

Following Deaton and Muellbauer (1981a), additive separability for a given partition on consumption $(Q_1,\ldots, Q_N)$ implies that for two different commodities i and j, such that $q_i \in Q_G$ and $q_j \in Q_H$, there exists an scalar λ which introduces the following restriction on the compensated price elasticities e_{ij}^*:

[1.2]

$$e_{ij}^* = \lambda e_i e_j w_j \quad \forall i \in G \text{ and } \forall j \in H,$$

where e_i and e_j are the expenditure elasticities of good i and j respectively and w_j is the budget share of good j. This expression collects the possibilities of substitutability between commodities belonging to different groups since the scalar λ is the only contact between them. Notice that this scalar does not depend on the groups to which i and j belong. Therefore, this restriction rules out any particular relationship between two different commodities or even groups. This is certainly a strong restriction when analyzing demand patterns specially at the individual level.

b) Weak separability.

Preference ordering is weakly separable if consumers group goods in a partition $(Q_1 ,\ldots, Q_N)$ such that, commodities belonging to an specific group are ordered independently from commodities outside that group, but groups are not necessary independent. Thus, the marginal rate of substitution among commodities included in the same group is independent of any other outside that group. According to this concept, the necessary and sufficient condition that characterizes a weakly separable utility function may be written as:

[1.3]

$$V(Q_1,\ldots,Q_N)=F(V_1(Q_1) ,\ldots, V_N(Q_N)),$$

being F an increasing function in all the arguments.

Following once more Deaton and Muellbauer (1981a), weak separability on a given partition $(Q_1, \ldots, Q_N)$ implies that for two different commodities i and j, such that $q_i \in Q_G$ and $q_j \in Q_H$, there exists an scalar λ_{GH} which introduces the following restriction on the compensated price elasticities e_{ij}*:

[1.4]

$$e_{ij}^* = \lambda_{GH} e_i e_j w_j \quad \forall i \in G \; and \; \forall j \in H,$$

where e_i and e_j are the expenditure elasticities of good i and j respectively and w_j is the budget share of good j. This expression collects the referred possibilities of substitutability between commodities belonging to different groups since the scalar λ_{GH} is the only contact between them. Nevertheless, this parameter depends on the groups we consider and hence it implies an increase in the number of price and expenditure responses relative to additive separability.

For empirical purposes, weak separability is a prerequisite for two-stage budgeting (Gorman, 1981). The first stage of this procedure implies to model the assignment of broad expenditure aggregates on different goods as functions of prices of those aggregates and income. Usually, this income allocation is not specified. As a second step, individual demand goods are modeled as functions depending on variables involved on that stage, that is, prices and total expenditure on those goods. In fact, if some goods belong to a separable subutility function, we can derive indirect utility and cost functions for the subgroups. Therefore, we can obtain demand functions for all the different commodities q_i included in the different groups Q_G, such that $q_i = g_{Gi}(x_G, p_G)$, being x_G total expenditure on group G and p_G the vector of prices of the commodities included in that group. In the opposite direction, the existence of specific demand functions defined on prices and on group expenditure require for weak separability. Notice that the main advantage of invoking two-stage budgeting is the reduction of the whole problem to a sequence of decisions.

c) Implicit or quasi-separability.

If preferences are defined from a cost function C instead of an utility function, we may characterize implicit separability by writing:

[1.5]

$$C(U,\ p) = Q(U,\ C_1(U,p_1),\dots,\ C_N(U,p_N)),$$

where each subfunction is increasing in utility U and p_G, being p_G a price index for group G.

From this type of separability, the implied restriction upon compensated elasticities may be formulated as:

[1.6]

$$e_{ij}^{*} = \mu_{GH} w_j \quad \forall i \in G \, and \, \forall j \in H.$$

Notice that the parameter μ_{GH} relates commodities belonging to different groups, and hence, relations among commodities belonging to different groups are restricted to group relations. Nonetheless, these parameters depend on the considered groups as in weak separability.

Implicit separability provides also a form of two-stage budgeting. The macrocost function C may be used to allocate expenditure among broad groups through the derivative property in log terms:

[1.7]

$$w_G = \frac{\partial \log C}{\partial \log C_G},$$

where w_G is the group G budget share. This characterizes the first step of two stage budgeting. We may also define the same partial derivatives in log terms for the bottom step and characterize the allocation within groups. Hence, individual budget shares will have the form:

[1.8]

$$w_{iG} = \frac{\partial \log C_G}{\partial \log p_i}.$$

d) Intertemporal separability.

Decisions upon consumption can be considered in an intertemporal frame. Therefore, additive, weak and implicit separability defined above within a period, that is in a static context, have an immediate translation when consumption decisions are considered in an intertemporal allocation process.

It is interesting to mention the implications for empirical purposes that both weak and implicit separability may introduce in an intertemporal framework. Relating commodity demands to current prices and expenditure, we are assuming intertemporal weak separability. If instead, preferences are implicitly intertemporally separable, commodity budget shares are functions of intertemporal utility and p. Demand analysis under this assumption requires the introduction of lifetime expected wealth and future anticipated prices into demand systems (Deaton and Muellbauer, 1981a). Desirable properties for demand systems such as homogeneity and symmetry would not apply in this context (Deaton, 1978).

Notice that these types of separability assumptions imply differences on how decisions are supposed to interact. In fact, modeling preferences one way or another we are restricting the possible substitution effects between goods within a period or among goods belonging to different periods.

In a static context, most of empirical works have focused their attention in modeling individual good demands, usually assuming weak separability between an specific good and the rest.[1] Implicitly, these studies are also assuming some sort of separability on preferences between consumption and leisure as well as intertemporal separability. Focusing the attention in a single good has the advantage of allowing to use functional forms so that specific explanatory variables may be introduced in a non-restricted way.

---

[1]See Atkinson, Gomulka and Stern (1989) and Atkinson, Gomulka and Stern (1990) for the demand on alcohol and tobacco respectively.

The immediate following step would be the simultaneous modeling of several related goods, usually assuming weak separability between them and the rest of goods.[2] Nevertheless, an important econometric problem, such as the group expenditure endogeneity, might arise from the conditional specification of demand goods as functions of prices on those goods and aggregated expenditure (Lafrance, 1991). Moreover, separability among non-durable goods has been usually tested and rejected. From this empirical regularity, it seems interesting to move to the consideration of complete demand systems. Accounting for the whole allocation problem, we can obtain unconditional demand elasticities suitable for policy and welfare analysis.[3]

Notice that an important problem that arises at this point is how to choose a suitable division for total expenditure. Some iterative procedures suggest to seek for a partition that minimize some distance between compensated elasticities obtained without separability assumptions and those which verify the implied separability restrictions among the elasticity parameters (Pudney, 1981).[4]

The problem may be formulated in the opposite direction. Instead of searching for a partition for total expenditure, we may be concerned about the criteria when grouping individual commodities. The composite commodity theorem (Leontieff, 1936) assesses that a group of commodities can be treated as a single one if prices display the same behavior. Nevertheless, this theorem has a limited translation to empirical analysis since relative price series are highly correlated. Moreover, relative prices evolve independently from demand patterns specially in the long run.

The analysis of complete demand systems on several goods defined from aggregation of single commodities, without separability assumptions, requires that all relative

---

[2]See Heien and Roheim (1990) for an empirical study upon demand for dairy products.

[3]For an estimation of a Linear Expenditure System upon total expenditure and different simulations of Value-added Tax reforms see Baccouche and Laisney (1991).

[4]Baccouche and Laisney (1991) suggest another method which is independent of the chosen initial partition.

prices must enter each demand function. Nevertheless, several empirical works detect a non-significance on most of the derived price effects due to a high correlation between individual commodity prices. It seems natural to seek for a procedure for grouping goods that reduce the number of price parameters to be estimated and hence that help to increase their significance. The concept of latent separability (Blundell and Robin, 1997) provides a procedure for commodity grouping by setting some restrictions on preferences that relax the mutual exclusivity of commodities implied when defining aggregated goods. The application of this concept implies a reduction on the number of aggregated commodities. Each of them is defined from a so-called exclusive good and participations on non-exclusive goods. This distinction allows us to test directly weak separability. If non-exclusive goods enter significantly in the wider aggregates, weak separability is rejected. It is worth to mention that following this procedure, the number of latent separable groups is perfectly defined by a rank test (Robin and Smith, 1994) but the method for choosing which goods must be characterized as exclusive is rather subjective.

Preferences may also be modified in such a way that decisions among periods may be related. Relaxing the usual invoked premise when formulating static preferences of intertemporal separability, we move towards a dynamic perspective on consumer's choices which implies an intertemporal planning over the life-cycle. There are some alternatives for preference specification such that past decisions may affect current utility. A common approach is based on the assumption that consumers display a habit behavior.[5] Hence, choices depend on tastes and these are constructed from past decisions. This fact translates into a higher correlation between current and past consumption rather than between the former and income. In fact, life-cycle hypothesis precludes from the possible dependence of consumption to income. This evidence, usually found in empirical works, is justified in terms of the presence of liquidity

[5]See Pollak (1970) or Spinnewyn (1981) for different intertemporal preferences that may capture habit behavior.

constraints.[6] Nevertheless, this correlation may vanish if preferences are modified allowing for both habits or durability in consumption (Blinder and Deaton, 1985).

If consumers do not behave myopically and anticipate the effects of today's decisions upon future consumption, we must allow past consumption to affect the marginal utility of current and future consumption. This assumption introduces dependence of current utility on current and past consumption. Such an structure on preferences relaxes the intertemporal separability assumption, which in fact, comes up as a testable proposition (Meghir and Weber, 1996).

1.2 - MOTIVATION

The previous theoretical background presentation leads us to the main points of research in this thesis. Preference specifications with separability assumptions enable the simplification on the substitution relationships between consumption goods. Simplifications on the relationships may be on both within a period or among different period's related consumed goods. The researcher has usually assumed different separability assumptions depending on the purpose and target of the study. Certainly, these premises simplify very much the formalizing of consumer's problem since preferences appear in more manageable ways. Our purpose along this dissertation is to test in different ways and contexts some specific invoked assumptions.

Another source of motivation comes from an interest in exploiting data at the microeconomic level. Applications of consumer theory have traditionally focused the attention in analyzing aggregated consumption and expenditure. The shortage in data bases at the microeconomic level as well as the high computational costs derived from their use have limited very much the possibilities of analysis using that type of data. In fact, most of the studies at the household level have been conditional on the

[6]See Hall and Miskin (1982), Zeldes (1989) and Runkle (1991).

availability of data. Both problems have been overcome in the last years. First of all, the increasing number of surveys conducted at the individual level has raised interest towards single units of decision. Moreover, computation using a large number of observations is less costly and allows to manage easily an important number of observations.

Studies at the microeconomic level use data which may validate directly the theory. In fact, the main advantage of working with household data is that we avoid the bias derived from aggregation across households. In this case, the level of the bias will depend on the non-linearity of Engel curves or demand equations. Furthermore, this aggregation process when constructing a representative agent rules out any possibility of heterogeneity. Opposite, staying at the individual level, consumers behave in an heterogeneous way. Thus, we are dealing with socioeconomic differences among households which must be controlled in order to derive coherent implications and consequences. Some of these characteristics are directly observed and measured but some are not. In fact, working at this level, it is rather difficult to justify models which do not account for unobservable individual heterogeneity (Deaton, 1992).

Most surveys at the household level present a cross-section design. The Encuesta de Presupuestos Familiares (EPF) in Spain is a good example. Nevertheless, cross-section data do not offer the possibility to analyze dynamic structures of consumption. Some cross-section surveys, such as the Family Expenditure Survey (FES) in the UK, are carried out along several years and offer the possibility to construct cohorts, but only a few follow household units for more than one period. The main advantages of using individual data with a panel structure are the allowance to incorporate dynamics and the possibility to control for both observed and unobserved heterogeneity.

A common important limitation for all the surveys is its specific design. In fact, surveys usually focus the attention in particular issues such as income or food consumption or even distribution of expenditure. None proposes an integral following of consumption, savings, labor and income variables. For instance, the Panel Study of Income Dynamics (PSID) reports information on income and food consumption since 1968 for the US. Also for the US, the Consumer Expenditure Survey (CEX) presents a panel design since interviewed households are asked on expenditure and income along 5 periods, although only 4 are usable. The short following-up of households is the main restriction of this survey. The Encuesta Contínua de Presupuestos Familiares (ECPF), conducted in Spain by the Instituto Nacional de Estadística (INE, 1985) since 1985, is specially designed to interview households across quarters up to 2 years on expenditure distribution, income and household characteristics. Unfortunately, most of the families do not stay the complete period and an important attrition ratio is present (see López, 1994). Nevertheless, it is possible to construct large enough panels with enough sample variation as well as a large enough profile for each household.

1.3 - CONTENTS OF THE DISSERTATION AND MAIN CONTRIBUTIONS

In this thesis, we use the ECPF survey covering the period 1985-1991. Selection and treatment of the different samples according to different objectives is described for each chapter. Nevertheless, some characteristics are common for the different data treatment. It is commonly assumed that all households face the same prices. Thus, explanations for household consumption and demand behavior are mainly captured throughout differences in expenditure and in family characteristics. Moreover, we control for observed heterogeneity but also for the unobserved effects.

The ECPF survey has not been exploited yet, to our knowledge, in a non-aggregated expenditure or consumption analysis considering its panel structure, and hence, controlling for unobservable heterogeneity. It has been used in some applied works in a somehow limited way, taking for instance samples with a cross-section design on demand systems (Labeaga and López, 1996). Some other studies have focused the analysis on specific goods.[7] Others have considered consumption as an aggregate taking profit of the dynamic possibilities the survey offers (Lopez-Salido, 1993, 1995). Therefore, this thesis constitutes in itself a contribution to the Spanish demand literature since it exploits for the first time the panel structure of this survey on a non-aggregated static demand framework and also on a non-aggregated dynamic consumption analysis.

In chapter 2, we specify and estimate a complete demand system using individual panel data. In particular, we specify and estimate an AIM (Almost Ideal Model, Deaton and Muellbauer, 1981b). Such an specification is also extended to a quadratic modeling in a restricted way just to detect up to what point our data fit in a rank two specification. The model includes socioeconomic variables which collect observed heterogeneity. Our purpose is to analyze the presence and effects of non-observed heterogeneity in the behavior of family units, specifically on income and price elasticities. Moreover, we control for the presence of another source of bias. The distinction between desired consumption and observed expenditure helps justifying the presence of zero record expenditures (Keen, 1986). In fact, we may associate zero record expenditures exclusively to infrequency of purchase since the sample is selected and treated according to a participation criteria on the analyzed goods. The applied procedure for the sample design requires at least a non-zero observation throughout the recording period. However, we allow that observed real expenditure does not coincide with desired consumption due to other circumstances different than infrequency.

From an initial well behaved static model, in chapter 3 we relax the usually invoked assumption of weak separability among goods. As pointed out, this assumption implies that the marginal rate of substitution among goods belonging to the same aggregate is independent from any other good included in another aggregate. Therefore, substitutability and complementarity relationships among individual goods

[7]See Labeaga (1992) and Labeaga and López (1997) for studies on the demand on tobacco and petrol

are reduced to relations between groups. This assumption has been usually tested and rejected. Anyway, a very disaggregated analysis of different goods is excessively costly in computational terms. Hence, it is interesting to seek an alternative procedure when grouping goods. We recall the concept of latent separability (Blundell and Robin, 1997). This term allows the distinction among exclusive and non-exclusive goods. The former enter only one aggregate whereas the latter may enter some composite aggregate goods at the same time. We assess evidence on the sensitivity of income and price elasticities upon the chosen exclusive goods. In fact, the construction of these pseudo-aggregates turns out to be subjective and hence its validity is questionable. Nevertheless, it allows to test directly weak separability among goods within the same period. If non-exclusive goods enter significantly in the construction of the aggregates, we will reject weak separability.

In this case, we estimate an Iterated Quadratic Almost Ideal Model (IQAIM), and hence, it is a non-restricted extension of the quadratic estimated version in chapter 2. We do not control for unobserved heterogeneity since the system is estimated in levels. Nevertheless, this is the first non-restricted rank three demand system for the Spanish economy.

Dealing with a panel data we have the chance to contrast weak separability among goods in different periods. Intertemporal weak separability implies that the assignment of expenditure within a period is independent of the assignment of life-cycle expenditure. In chapter 4, we assume dependence of consumer choices on tastes and the fact that these are constructed from past decisions. This dependence is described as a persistence of habits or inertia. In order to test this assumption, we model consumption in a life cycle framework, on four non-durable aggregates, in such a way that the utility function accommodates lags of consumption. Hence, we consider current utility as a function of current and lagged consumption. Such a model allows the inclusion of dynamics and hence to check intertemporal weak separability. Before

---

respectively and both for the Spanish economy.

doing so, we test for excess sensitivity of each category of consumption to income in order to ensure for consistency of the life-cycle framework.

Chapter 4 is based on a previous work by Meghir and Weber (1997). The purpose of their paper is rather different than ours. In fact, their target is to test for the presence of liquidity constraints by comparing two preference specifications set upon three aggregates of non-durable consumption. The first is derived from the marginal rate of substitution among goods, whereas the second comes from the Euler equations of a dynamic modeling. Under the null of no liquidity constraints both forms should represent the same sort of preferences. In their paper, both specifications are estimated using an homogeneous sample drawn from the CEX survey. As pointed out, this survey offers a very short temporal profile for each household, and in fact, this turns out to be crucial when testing intertemporal separability. Our main contribution in this chapter is that we present evidence on the importance of controlling for unobserved heterogeneity when testing for intertemporal separability on disaggregated consumption categories.

1.4 - MAIN CONCLUSIONS AND FUTURE RESEARCH

In chapter 2, we focus our attention in analyzing demand patterns on non-durable goods. The obtained results for rank two and rank three consistent specifications confirm the intuition about whether goods are necessities or luxuries. Although some of the price estimates are not relevant, they present the expected sign and move within a reasonable size. We also assess the importance of controlling for individual effects, both observable and unobservable, as well as for the errors derived from infrequency of purchase, when estimating price and income elasticities. To control for these different sources of errors leads to an specification in first differences and an estimation with IV, using lags of first differences of expenditure as instruments for expenditure. From this consistent estimation, we present different estimations and specifications in levels and in first differences controlling for each specific source of bias and we identify its effects on the derived income and price elasticities.

To ensure consistency on estimated income and price elasticities is crucial specially for simulations on optimal taxation. We provide evidence on the sensitivity on these parameters when controlling for individual effects and infrequency of purchase. Nonetheless, we are aware that the differences on the magnitude between consistent or biased parameters come up as relevant when simulating different tax changes. In order to evaluate the importance of the obtained results, we think that simulations on tax reforms constitute an important point of interest for future research.

As pointed out, this thesis contributes to the Spanish demand literature at the microeconomic level. It exploits individual data drawn from the ECPF covering the period 1985-1991. This period is characterized by a relative small price variation. This fact translates into a little significance of price effects due to multicollinearity. It seems interesting to extend our analysis by enlarging the sample such that we cover a wider period. By doing so, we expect a higher price variation and therefore a higher significance on price parameters.

Chapter 3 is specifically devoted to test on the usefulness of the latent separability concept. We derive income and price elasticities from a QUAIDS and observe that results under weak separability produce a better fit than those assuming latent separability. Moreover, the sensitivity of these parameters on the chosen exclusive goods leads to question the validity of this approach as a procedure for commodity grouping. As a by product, we test and reject weak separability. In fact, we observe that most of non-exclusive goods enter significantly in the construction of the pseudo-aggregates, and therefore, we can not assume weak separability among the exclusive goods. Finally, we compare the effects of price changes by simulations on revenue figures on weak and latent separable parameters and do not detect significant differences. These results suggest that more research has to be done in order to determine whether latent separability is a theoretical construction or has practical implications. As proposed for chapter 2, it seems also interesting to extend this analysis by enlarging the covered temporal profile so that it allows a better identification of price effects.

The analysis of consumption patterns in chapter 4, allows to test for intertemporal separability. As pointed out, we must test for excess sensitivity of consumption to income on the different consumption categories before recalling for a life-cycle modeling. For this first analysis, we detect that excess sensitivity vanishes on all the analyzed categories of consumption except on semidurables when including lags of consumption as explanatory variables. This result restricts the validity of the usual construction of non-durable aggregated consumption which includes semidurables.

We specify a similar problem than that formulated by Meghir and Weber (1997). The derived specification involves four observations per household, and therefore, using a sample derived from the CEX survey implies its treatment as a cross-section. Their estimations lead to a non-rejection of the intertemporal weak separability hypothesis. However, our specification is estimated using the ECPF and hence, we have a long enough profile for each household such that allows to control for unobserved heterogeneity. Also, we perform the same estimation upon a restricted profile for each household such that treats data also as a cross-section. We detect that intertemporal weak separability is rejected on the former whereas it is not upon the latter. A similar result is obtained by López-Salido (1994) whom rejects weak separability on an aggregated non-durable consumption analysis.

The problem presented in chapter 4 allows to characterize consumption on the analyzed goods according to a habit behavior. These results fall within the expected for all the consumption aggregates we construct except once more for semidurables. In this case we might expect a durable behavior. This non intuitive result is partly explained from the fact that we detect excess sensitivity of consumption on semidurables on income. We also produce evidence on the degree of intertemporal substitution for the analyzed categories by presenting the Elasticity of Intertemporal Substitution. Furthermore, we match higher values on these parameters with luxuries and lower values with necessities.

The modeling of preferences linking consumption between goods may be an issue of importance. In our problem, the set of specified preferences assume additive separability between consumed goods. This is certainly a restrictive hypothesis. Although we introduce consumption on other goods in a non-conditional way, we think that the modeling of preferences relaxing this assumption might be a possible extension on this analysis.

REFERENCES

ATKINSON, A.B., GOMULKA, J. and STERN, N.H. (1989): "Spending on alcohol: Evidence from the Family Expenditure Survey 1970-1983". *Economic Journal*,100, 808-827.

ATKINSON, A.B., GOMULKA, J. and STERN, N.H. (1990): "Household expenditure on tobacco 1970-1983: Evidence from the Family Expenditure Survey". ESRC Programme on Taxation, Incentives and the Distribution of Income, DP 134.

BACCOUCHE, R. and LAISNEY, F.(1991): "Describing the separability properties of empirical demand systems". *Journal of Applied Econometrics*, 6, 181-206.

BLINDER, A. and DEATON, A. (1985): "The Time-Series Consumption Revisited". *Brookings Papers on Economic Activity*, pp 465-521.

BLUNDELL, R., and ROBIN, J.M. (1997): " Latent separability: grouping goods without weak separability". Mimeo. *University College London.*

DEATON, A. (1978): "Specification and testing in applied demand analysis". *Economic Journal*, 88, 524-536.

DEATON, A. (1992). *Understanding Consumption.* University of Oxford. Clarendon Press.

DEATON, A. and MUELLBAUER, J. (1980a). *Economics and Consumer Behaviour.* Cambridge University Press.

DEATON, A. and MUELLBAUER, J. (1980b): "An almost ideal demand system". *American Economic Review*, 70, 312-26.

GORMAN, W.M. (1981): "Some Engel curves". In A.S. Deaton ed. *Essays in the theory and measurement of consumer behaviour.* Cambridge University Press.

HALL, R.E. and MISHKIN, F. (1982): "The Sensitivity of Consumption to Transitory Income. Estimates from Panel Data on Households". *Econometrica*, 50, 461-482.

HEIEN, D. and ROHEIM WESSELLS, C. (1990): "Demand systems estimation with microdata: A censored regression approach". *Journal of Business and Economic Statistics*, 8, 3, 365-371.

INE (1985): *Encuesta Continua de Presupuestos Familiares. Metodología.* Instituto Nacional de Estadística. Madrid.

KEEN, M. (1986): "Zero expenditures and the estimation of Engel Curves". *Journal of Applied Econometrics*, 1, 277-86.

LABEAGA, J.M. (1992): "A dynamic panel data model with limited dependent variables: an application to the demand for tobacco". Documento de Trabajo, 92/01, *UNED*.

LABEAGA, J.M. and LOPEZ, A. (1996): "Flexible demand system estimation and the Revenue and Welfare Effects of the 1995 Vat Reform on Spanish households". *Revista Española de Economía*, 13, 181-197.

LABEAGA, J.M. and LOPEZ, A. (1997): "A study of petrol consumption using Spanish panel data", *Applied Economics*, 29, 795-802.

LAFRANCE, J. (1991). "When is expenditure exogenous in separable demand models?". *West. Journal of Agricultural Economy*, 16, 49-62.

LEONTIEFF, W., (1936): "Composite commodities and the problem of index numbers". *Econometrica*, 4, 39-59.

LOPEZ, E. (1986): "La estructura del consumo en España en 1981. Una aplicación del modelo lineal de gasto". *Cuadernos de Economía*, 14, 89-106.

LOPEZ, A. (1994): "An assessment of the Encuesta Contínua de Presupuestos Familiares (1985-89) as a source of information for applied research". Working Paper, 53, *Universitat Pompeu Fabra.*

LOPEZ-SALIDO, J.D. (1993): "Consumo y ciclo vital: resultados para España con datos de panel". *Investigaciones Económicas*, 17, 285-312.

LOPEZ-SALIDO, J.D. (1995): "Time Non-Separabilities in Preferences: A Household Data Analysis". Working paper 9513 *CEMFI.*

MEGHIR, C. and WEBER, G. (1996): "Intertemporal Non-Separability or Borrowing Restrictions? A dissagregate analysis using the US CEX panel". *Econometrica*, 64, 1151-1181.

POLLAK, R.A. (1970): "Habit Formation and Dynamic Demand Functions". *Journal of Political Economy*, 78, 745-763.

PUDNEY, (1981): "An empirical method of approximating the separable structure of consumer preferences". *Review of Economic Studies*, 48, 561-579.

ROBIN, J.M. and SMITH, R. (1994). "Tests of Rank". Mimeo. *CREST Paris*.

RUNKLE, D.E. (1991): "Liquidity Constraints and the Permanent Income Hypothesis: Evidence from Panel Data". *Journal of Monetary Economics*, 27, 73-98.

SPINNEWYN, F. (1981): "Rational Habit Formation". *European Economic Review*, 15, 91-109.

ZELDES, S. (1989): "Consumption and Liquidity Constraints: An Empirical Investigation". *Journal of Political Economy*, 97, 305-346.

CHAPTER 2 - INDIVIDUAL HETEROGENEITY BIAS IN A DEMAND SYSTEM: AN ANALYSIS FOR THE SPANISH ECONOMY

2.1 - INTRODUCTION

The increasing number of surveys on individual data, collecting information on consumption and also including socioeconomic household variables, has raised the interest for microeconomic consumer behavior. Papers as Atkinson, Gomulka and Stern (1990) or Blundell, Pashardes and Weber (1993) among others, show the relevance of using microeconomic data to approach the analysis of consumer demand. The advantages of using disaggregated data are mainly that we avoid the problem of aggregation and therefore its implied bias. Also, we have a rich and wide sample from a statistical point of view. In the opposite direction, this sort of information is associated to measurement errors as well as zero expenditure records. This problem becomes more important the more disaggregated are the categories of expenditure that we analyze. Nevertheless, considering data at the household level, we can focus on idiosyncratic measurement errors, namely, those that are derived from tendencies of specific households which do not vary across time.

The shortage of data has conditioned very much the empirical work in demand analysis. Most of the papers exploit cross-section data bases, either on specific demands or complete demand systems. For the Spanish economy, the first estimate of a complete demand system using cohort data is Abadía (1984). López (1986) analyzes the structure of consumption in Spain using a cross-section while Lorenzo (1987) specifies and estimates different demand systems also from cohort data. Finally, Labeaga and López (1996) estimate a demand system on a pool of two microdata surveys from which some tax reforms are simulated. Nonetheless, most of the studies concentrate their interest on single goods using cross-section data.[8]

[8]See Matas (1991) for the demand of transport, García and Labeaga (1996) for the demand of tobacco and Moltó, Reig and Uriel (1990) for the demand of food.

The almost exclusive focus on cross-section data could be explained by the limited availability of panel data bases. Moreover, handling panel data is very costly from a computational point of view. However, the advantages derived from the use of this sort of data are clear (Hsiao, 1986). First, we can control for time invariant individual effects. Second, it allows us to introduce dynamics in the specification. For the Spanish economy the works using this sort of data are relatively scarce.[9]

The main goal of this chapter is to assess the importance of controlling individual unobservable effects and error measurement due to infrequency of purchases when analyzing demand patterns at the microeconomic level. Specifically, we produce empirical evidence on the demand for non-durable goods in Spain, using a panel of households derived from the Encuenta Contínua de Presupuestos Familiares (ECPF) survey. The functional form we adopt is based upon the Almost Ideal Model (AIM, Deaton and Muellbauer, 1980a, 1980b). Hence, we are able to analyze how budget and price changes affect household behavior. Also, we describe up to what point heterogeneity among consumers affects inferences on expenditure and price effects. Consistent and efficient estimates on these parameters will also be compared to previous results for the Spanish economy.

We concentrate the analysis on the specification and estimation of a demand system for 9 aggregated categories of non-durable expenditure. The analysis of decisions on those goods is assumed to be independent of decisions on durables and leisure, following the two-stage budgeting approach (Gorman, 1981). This procedure requires to invoke weak separability. Alternatively, it would be possible to model expenditure on non-durables conditional on durable decisions and on labor supply (Browning and Meghir, 1991) but our data does not contain information on either tenure of durable goods and variables

[9]See Monés, Salas and Ventura (1992) for saving decisions, and Collado (1993) and López-Salido (1993) and Cutanda (1995) for life-cycle models on consumption. See also Labeaga (1992) for the specification of a tobbaco demand equation. Finally, see also López (1998) for the demand for health care.

relative to labor supply such as the number of worked hours. In these circumstances, we are forced to invoke weak separability.

In a time series of cross-sections, an important part of the variation in the consumption pattern of the household will be due to individual effects both observable and unobservable, since the presence of heterogeneity among individuals is obvious. If this is the case, its control is a crucial part of the analysis. If this heterogeneity remains relatively constant over time, the panel structure allows us to control for it. Cross-section analysis cannot either control nor estimate these time invariant effects. The observable effects are measurable specific characteristics of the household, which are provided by the features of the family unit contained in the survey. Some of them refer to occupational status and proxy labor supply and their inclusion overcomes partly the restriction of the imposed separability between consumption and leisure. The individual unobservable effects are specific factors for the household units and are constant.[10]

The presence of zero expenditure records is quite common when working with dissagregated categories of expenditure. In this chapter, the constructed aggregates, as well as the treatment of the data, allows us to consider that all zeros can be associated to infrequency of purchase. This source of zero expenditures has usually been analyzed by introducing the distinction between non-observed desired consumption and observed expenditure (Keen, 1986). According to this difference, both variables are related through a policy of purchase for each household which depends on the purchase probabilities. Usually, these probabilities have been modeled as dependent on socioeconomic and demographic variables.[11] We also allow for the presence of errors in variables which are not explained in terms of infrequency purchases.

---

[10]This is a plausible assumption since the maximum period that a household reports its expenditure is two years.

[11]See Meghir and Robin (1992) for an example on a joint model for frequency of purchase and consumer demand.

The distinction between observed expenditure and desired consumption leads to a relation between observables plus an error from infrequency of purchases and an error in variables. Considering also the presence of unobservable heterogeneity, correlation between regressors and the error structure may arise from different sources. Nevertheless, given the dependence of the household policy of purchase on family characteristics and purchase habits, it seems reasonable to think that when controlling for those individual unobservable effects, we take into account, at least partly, the effects of infrequent purchases.

We shall present results for OLS and IV estimations in levels and first differences, controlling for individual effects and error measurement arising from infrequency of purchases and errors in variables. From the different estimations, we shall derive income and price elasticities. We cannot reject the presence of correlation between unobserved heterogeneity and the regressors that bias the pooled estimations. Moreover, using suitable instruments we are able to describe specific patterns for the bias derived from each source of error.

In section 2.2 of this chapter we characterize the theoretical framework in which the analysis is developed. The description of the sample, the treatment of infrequency and its association to the latent effects are analyzed in section 2.3. Section 2.4 is devoted to the econometric aspects. Results are presented in section 2.5. Finally, section 2.6 concludes.

2.2 - THE MODELING FRAMEWORK

2.2.1 - Separability assumptions

The analysis of consumer choices takes into account decisions between consumption and leisure, as well as the allocation of expenditure over all commodities. The study of the disaggregated categories of expenditure implies several complementary and substitutability relationships. In order to reduce them, consumer patterns are usually analyzed for broad groups. The logical approach is that consumption is partitioned into subsets that include commodities that are closer substitutes or complements among them. Weak separability has been the usual hypothesis in empirical demand analysis since it provides an approach for studying broad groups. According to this idea, the marginal rate of substitution among goods belonging to the same group is independent of any other good outside the group.

For a utility function V, this assumption allows to write the same preference ordering as:
[2.1]

$$V(c_1,\dots,c_n) = F(V_1(c_1),\dots,V_n(c_n)),$$

being V_1 ,..., V_n subutility functions, F some increasing function and c_i the consumption on good i.[12]

Weak separability is a pre-requisite for two-stage budgeting. According to the idea of two-stage budgeting, consumers proceed first to the allocation of total income among broad groups. In a second step, consumers decide how to distribute the group expenditure on individual goods. If a subset of goods appears in a separable subutility function, we can obtain demand functions for those goods as a function of expenditure on the group and prices of the different individual goods. In the opposite direction, the existence of a subgroup of individual demand functions depending only on prices of individual goods included in that group and on expenditure on the aggregate implies weak separability. The advantages of this approach are clear. Since it reduces the original problem to a sequence of decisions, each step requires only information on prices and expenditure on that specific decision level. Therefore, the maximization of the function V requires each c_i, depending on the n-prices and total income, to be the solution for the maximization of each V_i (Deaton and Muellbauer, 1980a).

[12]See Phlips (1974).

We concentrate exclusively on decisions over non-durable goods. Durable goods require specific models such as stock adjustment or probability of expenditure models. The analysis of these goods is out of the scope of this chapter. Considering only that set of goods we are also assuming separability between consumption and leisure. If leisure is weakly separable from consumption, decisions on leisure, and therefore on income, will be independent of the assignment of expenditure. This is a rather restrictive hypothesis. Browning and Meghir (1991) propose an alternative approach to overcome this problem. They model demand decisions conditional on hours and participation dummies which characterize labor supply. Even though we include some participation variables in a non restricted way, we can not model the suggested reduced form since our data does not contain information on worked hours.[13]

Besides, a temporal perspective requires that these decisions upon consumption and leisure must be taken considering not only present time but the next future. Intertemporal preferences would model complementarity and sustitutability between consumption and leisure in different periods. Here, we focus only on those relationships that come up from different categories of expenditure within the same period. We invoke for intertemporal weak separability on preferences so that the distribution of current consumption can be decided independently of the assignment of life-cycle expenditure.[14]

2.2.2 - The Almost Ideal Model (AIM)

We apply this formalization to the AIM (Deaton and Muellbauer, 1980a, 1980b). Three reasons may justify why we choose this specification for the demand functions. First, it is

[13]The number of worked hours may be proxied by introducing a dummy on participation since in Spain most of workers are full-time employed.

[14] The acceptance of this hypothesis depends very much on the analyzed good and on the period of time of expenditure we consider. See Marshall (1980) and Browning (1991) for aggregated consumption analysis. There is evidence also in microeconomic analysis which supports this hypothesis (Meghir and Weber, 1996). Nevertheless, in chapter IV we reject intertemporal weak separability for the broad consumption aggregates we construct.

a first order approximation to the demand functions that relates expenditure shares for each good with prices and expenditure with the form:

[2.2]

$$w_i = \alpha_i + \Sigma_j \gamma_{ij} \log p_j + \beta_i \log (\frac{X}{P}),$$

where P is a price index defined as:

[2.3]

$$\log P = \alpha_0 + \Sigma_k \alpha_k \log p_k + \frac{1}{2} \Sigma_k \Sigma_l \gamma_{kl} \log p_k \log p_l.$$

This functional form is almost linear except for the price index. Most of the empirical works approximate linearly this function using a Stone price index.[15] Since this price index enters all equations as a deflator for expenditure, we have a linear estimation problem. Also, the demand system for our problem is not constrained to expenditure or income but non-durable expenditure. Thus, the different analyzed goods form a separable group respect to the durable goods in the budget of the consumer. Our attention focuses in specifying the second step of a two step budgeting procedure. The first stage demand relation would yield non-durable expenditure as a function of a non-durable price, a durable price index and total expenditure.

The second advantage is that it offers the possibility to test the desirable properties for a demand system, namely, additivity, homogeneity which introduces within equation restrictions, and symmetry which introduces cross-equation restrictions on the parameters. The unrestricted estimation of the AIM is going to satisfy only additivity to keep full rank of the system. The rest of conditions will be testable in a simple way via parameter restrictions. Applied to the above model, these restrictions can be set as:

---

[15]We define $\log P_h = \Sigma_i w_{ih} \log p_i$, being w_{ih} the budget share of good i for household h. Pashardes (1992) shows that using such approximation, price effects estimates on the AIM may display a parameter bias specially if applied on individual data. Nevertheless, this bias depends mostly on the correlation between the expenditure parameters and the intercepts in the budget share equations. Estimations in first differences will overcome this inconvenient.

$$additivity: \quad \Sigma_k \alpha_k = 1, \quad \Sigma_k \gamma_{kj} = 0, \quad \Sigma_k \beta_k = 0,$$
$$homogeneity: \Sigma_k \gamma_{jk} = 0,$$
$$symmetry: \quad \gamma_{ji} = \gamma_{ij}.$$

Thirdly, this functional form is derived from a PIGLOG class of preferences that permit exact aggregation over consumers. These preferences are characterized by a cost function that in our case takes the form:

[2.4]

$$\log c(U,p) = a(p) + U\, b(p),$$
$$\text{being } a(p) = \alpha_0 + \Sigma_k \alpha_k \log p_k + \frac{1}{2} \Sigma_k \Sigma_l \gamma_{kl}^{*} \log p_k \log p_l,$$
$$\text{and } b(p) = \beta_0 \Pi_k p_k^{\beta_k}.$$

where α, β, and γ^* are parameters. Nevertheless, this flexible functional form represents a local approximation between different expenditure shares and prices and expenditure to an unknown general relation. The local character of this function implies a limitation of the significance and relevance of the performed test statistics. Global approximations with Fourier series overcome this problem. Nonetheless, symmetry is not testable since the functional forms include directly this property (Gallant, 1981).

Demographic and socioeconomic variables are usually introduced in the analysis of demand as explanatory factors of the behavior of the households. These variables are significant determinants of household consumption patterns. The usual procedure for including the household characteristics in the demand functions is demographic scaling (Pollak and Wales, 1981). This procedure implies reescaling the prices that enter the indirect utility function through a set of parameters, m_i, that depend on those variables. The quantities that would enter the direct utility function would be equivalent for typical units.

Given the indirect utility function depending on prices and income, $h_i\ (P,X)$, it can be reexpressed, reescaling with m_i as:

[2.5]

$$h_i\ (P,X) = m_i h_i\ (p_1 m_1,\ p_2 m_2, \ldots\ p_n m_n, X).$$

If we consider these arguments as the prices that enter the AIM, specifying:

[2.6]

$$log\ m_i = \Sigma_s\ \delta_{is} z_s,$$

demographic and labor status variables will enter linearly as regressors in the share equations (Pashardes and Baker, 1991).

Another important issue referred to the model is the rank of the demand system. Gorman (1981) demonstrates that the rank of the matrix of coefficients for the polynomial terms in income is at most three. Extending the AIM, which is initially rank two, to a rank three specification we obtain the following functional form:

[2.7]

$$w_i = \alpha_i + \Sigma_j\ \gamma_{ij}\ log\ p_j + \beta_i\ log\ (\frac{X}{P}) + \delta_i\ (log(\frac{X}{P}))^2.$$

This is the simplest quadratic extension of PIGLOG demands since the parameter of the quadratic term is a constant independent of *P*. Notice that integrability for a demand system with the above form requires $\delta_i = \beta_{i*}\ \varepsilon$ for all categories of expenditure.[16] Although this is a very simple extension, it will allow us to detect up to what point a rank two specification is too restrictive to impose on our data.[17]

---

[16]An integrable Quadratic Almost Ideal Demand System that did not verify such a condition might be formulated as follows:

$$w_i = \alpha_i + \Sigma_j\, \gamma_{ij}\ log\ p_j + \beta_i log\ (\frac{X}{f(p)}) + (\lambda_i\ /b(p))\ (log(\frac{X}{f(p)}))^2,$$

$$being\ b(p) = \prod_{i=1}^{N}\ p_k^{\beta_k}.$$

See Banks, Blundell and Lewbel (1997) for further details.

[17]Such a simple extension is regularly used in empirical analyses. See Blundell et al. (1993).

2.3 - SAMPLE DESIGN

2.3.1 - Description of the sample

The sample used in this chapter comes from the ECPF. This is a quarterly survey conducted by the Instituto Nacional de Estadística (INE, 1985) since 1985. The sample we work with covers the period 1985-1991. The survey established the interview of households throughout 8 quarters. Thus, the original design implied a rate of substitution of one/eight. Data analysis shows that there exists a higher level of attrition, which leads to a higher rate of substitution and fewer observations per household. Therefore, the actual sample is an unbalanced panel in the sense that we do not have the same number of temporal observations for each household. If families leave the survey according to an specific pattern, non-random attrition will imply biased and inconsistent estimates (Hausman and Wise, 1979). Nonetheless, representativity is preserved throughout all the considered period and hence, we may think that attrition is random.

When working with microeconomic data, we must deal with an important problem: zero expenditures on several categories of consumption will be usual. This constitutes an important justification for grouping goods.[18] Three reasons may generate zero expenditures: first, as a result of a corner solution; second, non-participation; and third, infrequency of purchase (Pudney, 1989). The nature of observed zeros depends very much on the category of expenditure we consider. As suggested by Blundell and Meghir (1987a), a suitable assumption is that there is only one source of zeros for each good. However, categories with zeros mainly due to non-participation require specifications that include the participation decision. A conditional demand system is the most suitable framework to model demands on these goods.[19] Since we are

[18]Another important justification for grouping goods is the reduction of the problem to a reasonable number of equations.

[19]See Lee and Pitt (1986) and Pollak (1969). These approaches are not feasible when analyzing more than 3 goods.

interested in evaluating the potential biases that infrequent purchases might generate, we restrict our attention to goods for which infrequency of purchase can be reasonably assumed to be the sole generator of zeros.[20]

We select 9 groups of expenditure on non-durables. These groups do not cover total expenditure on non-durables since we exclude consumption on those goods for which expenditure may be conditional on participation variables. The expenditure goods we consider are: food and non-alcoholic beverages, alcoholic beverages, clothing and footwear, rents and house keeping expenditures, fuel for housing, transport and communications, services and leisure, household non-durables and other non-durable goods.

The final sample we work has 4372 households observed throughout 6 quarters. It has been selected according to two criteria: first, we maintain those households that stay at least 6 quarters. A cohort analysis shows that none of the households that enters the survey in 1985 and the first 3 quarters of 1986, completes the 8 quarters. Selecting only those households that respond 8 quarters, we loose representativeness of an important period in terms of price variation. Moreover, most of the families stay in the survey 6 or less quarters. In order to have a balanced panel we drop the last two observations for those households who report 8 quarters and the last one for those that report 7. This temporal profile gives enough lags of the independent variables, which may be used as instruments. The second sample selection criteria is that we require consumption participation on the analyzed goods. Panel data allows the identification of zeros due to non-participation. We assume that a single non-zero expenditure observation throughout the observed period identifies the household as a consumer on that group. By doing so, we associate the remaining zero expenditures to infrequency of purchase. For food and non-alcoholic beverages we require all reported expenditures to be positive.

[20]The main expenditures on non-durable goods we are not including are petrol and tobacco.

2.3.2 - Infrequency of purchase

Infrequency zeros arise due to the indivisibility of expenditure in such a way that the specific moment of purchase does not fall within the monitoring period covered by the survey. Moreover, infrequency is due to indivisibility of goods and also to searching costs. In the absence of these costs (or highly storage costs) and perfect divisibility, consumers would distribute their expenditure in such a way that household desired consumption would coincide with observed expenditure whatever the period we considered.

The framework we consider to analyze infrequency of purchase is based on the differentiation between desired non-observed consumption c_{hk} and observed real expenditure e_{hk} for household h on good k (Kay, Keen and Morris, 1984 and Blundell and Meghir, 1987b). The stochastic relationship among both takes the form:

[2.8]

$$e_{hk} = d_{hk}\{c_{hk}\} \ / \ pr_{hk} + u_{hk},$$

where u_{hk} is a perturbation of expenditure; d_{hk} is a random variable distributed as a Bernouilli and pr_{hk}, is the probability of purchase during the interview period of good k by household h. Two error measurement sources come up from this distinction. The purchase non-purchase decision implies an error that explains itself the presence of zeros. The implied consequence is that real observed expenditures are biased estimators of desired consumption. The usual approach has been to instrument expenditure with income.[21] This procedure does not require the knowledge of the purchase probabilities. Also, we introduce the possibility that observed real expenditure does not coincide with desired consumption due to other circumstances different than infrequency. This element displays a random behavior. Considering this model, we take into account both the effects of the household purchase policy and other errors in variables which are not determined by infrequency purchases.

[21]See Keen (1986) for an application to linear Engel curves.

Meghir and Robin (1991) suggest a method to deal with unobserved consumption that takes into account the purchase probabilities and also allows to deal with non-linear models. The suggested procedure requires obtaining those probabilities of purchase over the whole sample, conditional on demographic characteristics. In a second step, those households that have positive expenditures are selected and for those units, desired consumption is constructed by reweighting observed expenditures with the obtained probabilities. Working with a panel data, this approach implies a high cost in terms of the number of observations we loose since the selection of positive expenditures will withdraw those households for whom continuous in time observations are not available.[22]

The distinction between observed expenditure and desired non-observed consumption according to equation [2.8] for the AIM. leads to a stochastic relation with the form:
[2.9]

$$\frac{d_{hk}\{p_{kt}q_{hkt}\} \,/\, pr_{hk} + u_{hkt}}{\Sigma_k d_{hk}\{p_{kt}q_{hkt}\} \,/\, pr_{hk} + \Sigma_k u_{hkt}} = \frac{p_{kt}q^*_{hkt}}{\Sigma_k p_{kt}q^*_{hkt}},$$

being $p_{kt}q_{hkt}^*$ and $p_{kt}q_{hkt}$ observed and desired expenditure for period t on good k by household h. As pointed out, the purchase policy is modeled through the probability of observing a positive expenditure and the dummy variable. Notice that we assume that both are time invariant since these probabilities are assumed to be dependent on household characteristics.[23] Hence, we model a time invariant purchase policy. Introducing this purchase policy into the demand equations, we obtain the following relation between observables:

[22]In our case, the number of periods we observe a household purchasing could be used as the purchase probability to apply the method proposed by Meghir and Robin. See Labeaga and López (1997) for an application on petrol consumption for the Spanish economy using a panel data.

[23]Each household is followed across 6 quarters. It seems reasonable to assume that family characteristics are steady during the observed period.

[2.10]

$$\frac{p_{kt}q^*_{hkt}}{\Sigma_k p_{kt}q^*_{hkt}} = \alpha_k + \Sigma\gamma_{kj} ln\ p_{jt} + \beta_k(ln\ [\Sigma_k \frac{p_{kt}q^*_{hkt}\ *\ pr_{hk}}{d_{hk}} - \Sigma_k u_{hkt}]) - \beta_k\ ln\ p_t$$

$$+ \alpha_{hk} + \varepsilon_{hkt} + \eta_{hkt},$$

$$being\ \ \eta_{hkt} = \frac{d_{hk}\{p_{kt}q_{hkt}\}\ /\ pr_{hk} + u_{hkt}}{\Sigma_k d_{hk}\{p_{kt}q_{hkt}\}\ /\ pr_{hk} + \Sigma_k\ u_{hkt}} - \frac{p_{kt}q_{hkt}}{\Sigma_k p_{kt}q_{hkt}}.$$

Share expenditure equations include the parameter α_{hk} which represents the individual unobserved effects. Notice also that we are considering errors in variables for expenditure on each individual category and for total non-durable expenditure both time variant. Moreover, the error structure has been derived accounting also for errors in variables in the left-hand-side, that is, in the budget shares. As far as the error in variables occurs in an specific good, no special problem comes up. However, since the denominator is total expenditure, if errors in variables are present, it has a non-polynomial structure and no obvious solution exists. We recall Hausman, Newey and Powell (1995). They do not find significant differences specifying the dependent variable in levels and in budget shares when estimating Engel curves. This empirical evidence is used to assess that the errors in variables in the denominator of the left-hand-side variable in a budget share specification do not create an important problem. Therefore, omitting this source of error and assuming a linear structure we can account for an instrumental variables estimation procedure and the above expression becomes:

[2.11]

$$\eta_{hkt} = \frac{d_{hk}\ [p_{kt}q_{hkt}\ ]\ /\ pr_{hk} + u_{hkt} - p_{kt}q_{hkt}}{\Sigma_k p_{kt}q_{hkt}}.$$

Appending this expression to the stochastic error, we describe the whole structure error with:

[2.12]

$$\eta_{hkt} + \varepsilon_{hkt} = \left(\frac{d_{hk}}{pr_{hk}} - 1\right) w_{hkt} + \frac{u_{hkt}}{\sum_k p_{kt} q_{hkt}} + \varepsilon_{hkt}.$$

The first component collects information about the purchase policy. In fact, it consists of an interaction between policy of purchase and desired budget shares. Notice that the latter depends on prices and total non-durable desired expenditure and hence, it varies individual and temporally. Besides, the probabilities of purchase are usually assumed to be dependent on household characteristics. Therefore, we may conclude that the error component related to the policy of purchase must affect all categories of expenditure the same way. The second error term refers to errors in variables different than infrequency. Finally, the third term corresponds to the usual stochastic error. Notice also that we do not consider in this error structure an only time dependent component since the used microeconomic panel data has a reduced temporal profile.

Correlation between this error structure and the regressors is obvious. The implied inconsistency may come up from correlation between total expenditure and infrequency and from the errors in variables component. As pointed out, the derived effect from infrequency of purchase must be the same on all expenditure categories. However, the direction of the implied bias from correlation between expenditure and errors in variables is not well determined. Additionally, we must add another source of bias derived from the correlation between individual unobserved effects and the regressors. Considering the different sources of inconsistency, we can predict that the direction of the total bias will not affect all categories of expenditure in the same direction. OLS estimates will be affected by all these sources of inconsistency but must verify additivity conditions as well, and hence, we expect that the bias will be compensated among categories.

In spite of the ambiguity of the total bias, we analyze the error term related to the purchase policy and conclude how probabilities of purchase may affect the sign and magnitude of the bias. Hence, focusing only in the purchase policy error term, the derived bias from OLS estimations has the following form:

[2.13]

$$plim \sum_n ln\left(\sum_k p_{kt} q_{hkt} / P_t\right)\left(\frac{d_{hk}}{pr_{hk}} - 1\right) w_{hkt} \Big/ \left[\sum_n ln\left(\sum_k p_{kt} q_{hkt} / P_t\right)\right]^2 .$$

which has a positive sign. Now, considering the partial derivative respect to the probability of purchase, we expect that an increase in that probability will lead to a lower bias. Obviously, we will expect that those categories with a lower incidence of zero records will display a lower bias derived from infrequency than those with a higher incidence of null records.

2.4 - ECONOMETRIC ISSUES

2.4.1 - Relative price variation

A usual problem in demand analysis at the microeconomic level is the lack of relative price variation. Although our sample covers a time-span of 7 years, relative prices evolve very much in line. Price series display a high correlation among categories and hence, we cannot significantly identify different price effects for all expenditure subgroups due to multi-collinearity. Moreover, all households face the same prices. It would be possible to obtain temporal series for different regions. This would introduce some additional variability. Nevertheless, variables related to spatial location of households are not provided by the INE.

2.4.2 - Specification and estimation

The main target of this chapter is to assess the importance of controlling individual effects, both observable and unobservable, for describing demand patterns and derive expenditure and price elasticities. We specify a static model that includes relative prices (to the omitted category of expenditure), real non-durable expenditure and quarterly dummy variables. The specification of prices relative to one category imposes the homogeneity restriction. However, the initial AIM relates linearly budget shares with prices and deflated income. Two step budgeting requires to use expenditure on the considered goods instead of income. By doing so, we are introducing endogeneity in the model and we will need to instrument expenditure. Recalling demographic rescaling and the distinction between desired and observed expenditure, we include in the AIM both observable and latent individual effects. According to the points considered above, the final expression for the specification of the demand equation is:

[2.14]

$$w_{hkt} = \alpha_k + \Sigma_{kj}\gamma_{kj} log\ \ p_{jt} / p_{Nt} + \beta_k log(\frac{X_{hkt}}{P_t}) + Z_h \lambda_k + \alpha_{hk} + \upsilon_{hkt} + \varepsilon_{hkt},$$

where α_{hk} captures unobserved heterogeneity and v_{hkt} is the error term related to infrequency of purchase.

Demand equations across goods have the same regressors and hence, estimation equation by equation will provide consistent but not efficient estimates. Although we have grouped goods so that direct complements and substitutes are included in the same category, correlation between different equations has to be considered.

There are several econometric techniques that, applied on panel data sets, control for unobserved heterogeneity among individuals. The treatment of these latent individual effects as fixed or random does not imply any gain in terms of specification. Working with samples with a wide cross-section variation, it is desirable to make unconditional inferences to the sample and therefore, to treat individual unobservable effects as random (Mundlak, 1978). This assumption implies that error terms will have a mixed structure. The GLS estimator (Balestra and Nerlove, 1966) is going to be consistent and efficient under the hypothesis of absence of correlation between regressors and errors. If such correlation did not exist, the GLS estimator is not consistent. Besides, these effects are assumed to be i.i.d. Homoskedasticity seems to be a rather restrictive hypothesis. Using budget shares as dependent variables we need to make a heteroskedasticity correction. Calculating standard errors with expressions that take into account the presence of heteroskedasticity of an unknown form, the estimators will also be robust.[24]

When working with individual data, it is quite usual to detect the presence of correlation between the error structure and the regressors. In our analysis, this correlation may arise, first of all, from the individual unobservable effects and expenditure since the former can be described as a function of the latter. Estimating equations in levels, we do not remove these individual unobservable effects and therefore we should detect correlation. The usual treatment for this problem is to instrument expenditure. An available straightforward instrument seems to be income, which is highly correlated with expenditure. Nonetheless, this instrument may also be correlated with unobservable heterogeneity. Notice that using income as instrument we do not take into account the invariant nature of the latent effects. Other possible invariant in time instruments refer to characteristics of the family, but they are usually included in the equation as observable individual effects and then there are identification problems (see Browning and Meghir, 1991). This problem may be overcome by using as instruments the individual means of those variables which are not correlated with the latent effect (Hausman and Taylor, 1981). In our case, we do not have any regressor uncorrelated with the effects which is variable across individuals and time. Moreover, correlation between the observable and unobservable individual effects is expected. For this sort of correlation we do not have any available instrument since these effects are time invariant. If they were not, first differences of the socioeconomic variables

[24]White (1980).

could be suitable instruments although sometimes they do not produce good results on some variables.

Finally, another possibility to deal with the presence of invariant in time individual random effects, which are correlated with the regressors, is to remove the individual effects, both observable and latent, by taking first differences. Estimating by OLS the equations in first differences, we will obtain consistent estimators, given the static nature of the specification and assuming exogeneity of expenditure. In fact, dealing with regressors linearly correlated with the latent effects, the optimal estimator (Minimum Distance or Maximum Likelihood estimator) coincides with the OLS estimator applied to equations in first differences (Chamberlain, 1982 and 1984 and Arellano and Bover, 1990). In fact, if we are dealing only with unobservable effects, the Within Groups procedure will provide consistent estimators.

The above analysis has only taken into account correlation between the individual latent effects and the regressors. We settled the distinction between unobserved consumption and observed expenditure. From this difference, we deduced the presence of time dependent errors derived from infrequency. We also considered the presence of errors in variables. Once more, the implied bias can be solved by instrumenting those variables from which correlation arises. Again, non-durable expenditure may be proxied with income. Nonetheless, income may be correlated with infrequency of purchase since probabilities of purchase depend majorly on specific household variables. First differences of non-durable expenditure lagged one period may be a suitable instrument for the estimation in levels if the infrequency errors are i.i.d. Equations in differences, again under the null of uncorrelated measurement errors, will display first order but not second order serial correlation. In this case, differences of expenditure lagged two periods will be orthogonal to the first differences errors. Note that if errors of measurement have an invariant nature, they will be dropped out in the first differences estimation and hence, we will not need to use instrumental variables techniques.

Income is the usual instrument either if there is correlation between individual effects and expenditure or in the presence of error measurement due to infrequency.[25] As pointed out, dependence of the probabilities of purchase on socioeconomic characteristics may translate into a correlation between income and infrequency. Moreover, it is worth to mention that income is going to be a meaningful instrument only if we accept weak separability between consumption and leisure. If this is the case, the decision on leisure, and therefore on income, can be considered exogenous related to the consumption on non-durable goods choice. Summers (1959) and Livitan (1961), among others, assume that income is uncorrelated with the error term associated to a linear Engel curve. Their assumption is based on the Friedman assessment that permanent income and transitory consumption are uncorrelated. Under the null of absence of correlation between income and the stochastic disturbance, a test of exogeneity of expenditure can be derived. Opposite, Attfield (1978) points out that this is a very strong assumption working with individual data. Household data will display a high correlation between income and specific unobservable effects. In this case, possible alternative instruments are lags or leads of expenditure.

For the rank three specification, we specify the simplified quadratic parameter linearized extension. Working in a non-linear context and in the presence of measurement errors, the IV procedure will not provide consistent estimates whatever set of instruments we use (Amemiya, 1985), Hausman, Newey and Powell, 1988). The reason has to be found in the fact that error measurement is non separable from the true variable. Nevertheless, if this error term is time invariant, first differentiation will cancel it and despite of the non separability structure, the observed regressor will not be correlated with the error structure in first differences. If it shows a time dependency, we can use Hausman, Newey and Powell (1995) repeated measurement procedure.[26]

[25]Income appears regularly as underestimated in data bases at the microeconomic level (see Raymond, Oliver and Pujolar, 1995) Expenditure may come up as misreported as well but this is already captured by the error measurement term. Nevertheless, underestimation of the former exceeds the latter, specially considering that most of the samples, including ours, are designed in order to study the structure of expenditure. Therefore, we must cast doubts about the adequacy of income as a proxy for expenditure.

[26]This technique proposes to use alternative variables to construct adjusted expenditure, and use it as instrument. Possible variables are education and age which proxy expenditure and will not be correlated with

2.5 - RESULTS

2.5.1 - Discussion upon rank two and rank three specifications

Estimations in levels, either considering data as a pool (OLS), or introducing the presence of random heterogeneity (GLS) generate certainly different results from those obtained from estimations in differences (either WG and first differences). This suggests that there are unobservable effects that bias the estimations in levels because these effects are correlated with the regressors.

We perform a Haussman test to detect more formally the presence of correlation between individual effects and regressors by comparing GLS and WG estimators. Notice however that the latter will only be consistent under the null of abscense of measurement errors.[27] Still, the non-correlation hypothesis is rejected for all the equations. Also, an F-test for the presence of individual effects leads us to reject the null of homogeneity of the individual effects (see Tables A.2.1 through A.2.8 in Appendix A.2 for both test results). Therefore, to treat data as a pool leads to biased estimators due to the omission of the individual unobservable effects. The presence of latent heterogeneity as well as infrequency of purchase errors will require to instrument expenditure.

First, we present results of estimations in levels. OLS estimations include as explanatory variables the socioeconomic characteristics of the household. The introduced variables refer to the labor situation and activity (dummies for the head of the household in non active, self-employed or unskilled situation). The number of members of the family, number of members under 14 years old and the number of earners are characteristics of the family. We also include 3 quarterly dummy variables to capture seasonality in

---

the non linear stochastic disturbance. Other possible instruments are lagged expenditure or even future expenditure. Under rational expectations, observations located in the following future will be independent of current information and will also proxy current consumption.

[27]If the error measurement component displays a time invariant structure, WG estimators will also be consistent.

consumption. Columns 1 and 2 in Tables A.2.1 through A.2.8 show the parameter results while Tables 2.1 and 2.2 report the derived price and expenditure elasticities.[28] First column picks up OLS results. Column 2 presents results using one lag first differentiated expenditure as instrument for non-durable expenditure. First differences of the mentioned characteristics of the household are also used as instruments.

As a general pattern for all estimations, most of the price parameters turn out to be non-significant, whereas the quarterly dummies are highly significant. Although the coeficients are non-significant, the own price elasticities derived from the values fall within the expected range and are significantly different from zero.[29] Moreover, these price effects are jointly significant for most of the categories (see tests on Tables A.2.1 to A.2.8). We are analyzing a period of 28 quarters; for this short period, the variation on relative prices is very small, and also, we detect a high correlation between the different price series. Partly, the variation of the relative prices might be associated to seasonality but this effect is already captured by the quarterly dummies. So, we are pretending to capture dissagregated price effects with a small variation. The parameters for these quarterly dummies characterize perfectly seasonality, specially on those categories that follow a different consumption pattern depending on the period of the year we consider. When we do not introduce these quarterly dummies, price effects are significant, but their sign and magnitude are counterintuitive.

---

[28]Price and expenditure elasticities are evaluated at the mean, for a representative consumer, according to the following expressions for the linear and quadratic specifications:

$$\overset{*}{\varepsilon}_{ii} = \frac{\gamma_{ii}}{\bar{w}_i} - 1, \qquad \varepsilon_i = \frac{\beta_i}{\bar{w}_i} + 1$$

and

$$\varepsilon_{i\,sq} = 1 + \frac{\beta_1 + 2\beta_2 ln(X/P)}{\bar{w}_i}.$$

[29]Standard errors are calculated by bootstrapping.

Expenditure parameters are estimated with precision. Their sign and magnitude are in accordance with a priori expectations. We also characterize those goods that typically come up as luxuries, namely, clothing, transport and communication and services.

The socioeconomic variables are highly significant in general for the OLS estimation. Significance of individual observable variables for the OLS estimation has to be interpreted as a proof for the presence of latent effects since both are highly correlated. The relevance of these variables for the IV estimation depends on the category of expenditure we analyze. Labor variables do not seem to affect very much in none of the categories. However, family composition comes up as very significant on all equations, specially on those that are characterized as luxuries.

Turning to estimations in first differences, column 3, in Tables A.2.1 to A.2.8, shows OLS estimates. Columns 4 and 5 are IV estimators using two period lags of differentiated expenditure and income respectively. Since first differentiation implies dropping out all the variables that do not change in time, we do not include the socioeconomic variables. We analyze changes in these variables along the 6 periods we consider (see Table A.1.1) and determine that most of the heads of household do not vary their characteristics. In particular, we determine a very low percentage of households that move from a given position or status to any other along the followed period.

Price effects, as above, are in general non-significant whereas expenditure effects are well defined. Again, a specification with price variables and without quarterly dummies raises very much the significance of the former, but instead, the derived elasticities present non-intuitive values since the parameters are biased. We can conclude that most of the relative price variation is due to seasonality. Both price and expenditure elasticities are significant and fall within the expected range characterizing as luxuries the same categories than the estimations in levels.

Tables 2.1 and 2.2 summarize price and expenditure elasticities for all estimations. Our reference starting point is an specification in first differences, using as instrument for total expenditure two lags of first differentiated expenditure (column 4). Such estimation controls for all sources of error. In fact, these will be suitable instruments under the null of i.i.d. errors. We test for the presence of correlation in the error structure and reject it for all categories (see Tables A.2.1 to A.2.8).

From this consistent set of parameters, we move to the same first differences specification but deriving OLS estimates (column 3). In this case, we control also for unobservable heterogeneity but we do not take into account the possible correlation between first differences of the infrequency error term and errors in variables and differences of total expenditure. A test on the comparison of both sets of instruments provides information about the possible correlation of expenditure and measurement errors and also on the endogeneity of expenditure. Notice first that the obtained results are very similar in both income and price effects. Anyway, the test-value we obtain is very low: 0.01 (96 d.f.).[30] From this evidence, we reject endogeneity of expenditure and also a possible correlation between differences of expenditure and measurement errors. However, intuition suggests that both variables might be correlated in level terms specially with the infrequency error term. Therefore, we think that measurement errors may not be time dependent and hence, they cancel in first differences specifications. The same conclusion can be derived if we compare both sets of estimations with WG parameters. Although not reported here, income effects do not differ at all from the previous commented results. Price effects fall within the same rank although are not that close. Notice anyway that for all estimations these parameters do not come up as significant as income effects.

We now move to a levels specification using one lag first differentiated expenditure as instrument for total expenditure (column 2). If the error measurement structure is i.i.d.,

---

[30]The Durbin test for the possible endogeneity of expenditure has the form:

$$(\beta_{olsq} - \beta_{iv})'(V_{iv} - V_{olsq})^{-1}(\beta_{olsq} - \beta_{iv}) \sim X^2(m).$$

orthogonality between the proposed instrument and the error components is ensured; hence, differences, if present, must be explained in terms of correlation between the individual effects and the regressors. We observe that these estimators in the levels equations are quite close to the estimators of the first differences equations, although we may characterize that the former are relatively downwardly biased except for housing and domestic fuel. Price effects follow the opposite direction. We conclude that correlation between expenditure and the latent effects tends to bias downwardly income parameters and opposite for price effects. If instead, we use current differentiated expenditure (not reported here either), we obtain very similar results. This supports that correlation between error measurement and differences of expenditure is not relevant. Once more, this supports that error measurement may display an invariant in time behavior.

The levels specifications without controlling for any source of error implies correlation between expenditure and individual effects and measurement errors. From the above result, we described the direction of the bias implied from the former source of correlation. We now compare consistent estimates on column 4 with OLS pooled estimates in column 1, taking into account also the parameters on the second column. We observe that the direction of the bias implied only from measurement errors depends on the category of expenditure. We detect an upward bias on housing, domestic fuel, services and house non-durables whereas a downward bias for the rest. In section 2.3.2, we characterized that the expected bias from the household policy of purchase tended to bias upward the parameters, especially on those categories with a higher incidence of zeros. This pattern is only followed by 4 of the analyzed categories from which only house non-durable expenditure is significatively affected by the presence of zero records. Hence, we conclude that there are also errors in variables, different than infrequency of purchase, included in the error measurement structure, which bias the results in the opposite direction.

Back to IV estimations in first differences, if instead of lagged differences of total expenditure, we use differences on income (column 5), we also control for unobservable heterogeneity, but we do not take into account either the possible correlation between infrequency and income. Nevertheless, from the previous results we deduce that measurement errors display an invariant in time nature and hence they cancel out in this specification in first differences. Comparing both OLS and IV estimators, using income as instrument for expenditure, all in first differences, we obtain again a test for endogeneity of expenditure. This test can also be reinterpreted as a test for the validity of income as instrument, that is, a test on separability between income and expenditure. We obtain a value of 4.96 which must be compared with a χ^2 with 96 d.f. This result implies a non-rejection of income as a suitable instrument. Nonetheless, if we perform the same test only upon the subset of expenditure parameters, this value raises up to 55.3 (8 d.f.) which clearly rejects the null. As pointed out, the hypothesis of orthogonality between the stochastic disturbance, ε_{ikt}, and income seems questionable when working with data at the individual level. Hence, we conclude that the obtained high value, when testing exogeneity of expenditure, is an evidence of non-separability between consumption and leisure.

Table 2.1 - EXPENDITURE ELASTICITIES

	1	2	3	4	5	6	7	8
food	0,652	0,687	0,693	0,699	0,629	0,696	0,698	0,700
	(0,011)	(0,050)	(0,016)	(0,023)	(0,063)	(0,013)	(0,018)	(0,021)
alcoholic bev.	0,825	0,992	1,023	1,053	0,861	1,021	1,044	1,17
	(0,081)	(0,366)	(0,163)	(0,151)	(0,491)	(0,135)	(0,143)	(0,205)
clothing	1,388	1,582	1,642	1,619	1,671	1,640	1,660	1,593
	(0,060)	(0,292)	(0,122)	(0,129)	(0,335)	(0,123)	(0,123)	(0,144)
housing	0,844	1,763	0,711	0,707	0,927	0,710	0,677	0,729
	(0,337)	(0,368)	(0,322)	(0,329)	(0,375)	(0,116)	(0,114)	(0,233)
fuel	0,480	0,374	0,349	0,357	0,234	0,353	0,333	0,377
	(0,180)	(0,290)	(0,117)	(0,116)	(0,313)	(0,114)	(0,115)	(0,236)
transp. and comunic.	1,407	1,599	1,616	1,673	1,383	1,593	1,599	1,491
	(0,329)	(0,528)	(0,247)	(0,253)	(0,567)	(0,249)	(0,235)	(0,277)
services	1,275	1,192	1,242	1,230	1,567	1,248	1,248	1,256
	(0,041)	(0,171)	(0,114)	(0,122)	(0,180)	(0,117)	(0,116)	(0,136)
house non-durables	1,027	0,946	0,947	0,959	0,621	0,949	0,958	0,975
	(0,017)	(0,110)	(0,019)	(0,026)	(0,106)	(0,015)	(0,015)	(0,052)

Note: Standard errors are in parentheses

Table 2.2 - OWN-PRICE ELASTICITIES

	1	2	3	4	5	6	7	8
food	-0,636	-0,651	-0,767	-0,769	-0,729	-0,585	-0,773	-0,949
	(0,011)	(0,013)	(0,019)	(0,019)	(0,020)	(0,015)	(0,0117)	(0,027)
alcoholic bev.	-1,772	-1,804	-1,652	-1,660	-1,603	-2,115	-1,756	-2,583
	(0,360)	(0,375)	(0,387)	(0,387)	(0,389)	(0,353)	(0,381)	(0,306)
clothing	-0,128	-0,159	-0,105	-0,115	-0,079	-1,291	-0,035	-0,984
	(0,136)	(0,139)	(0,154)	(0,153)	(0,155)	(0,184)	(0,160)	(0,196)
housing	-0,949	-0,962	-1,774	-1,780	-1,499	-1,075	-1,990	-2,096
	(0,110)	(0,114)	(0,121)	(0,129)	(0,134)	(0,100)	(0,115)	(0,177)
fuel	-0,668	-0,673	-0,988	-0,987	-1,006	-0,929	-0,986	-0,626
	(0,115)	(0,116)	(0,122)	(0,125)	(0,126)	(0,128)	(0,123)	(0,122)
transp. and comunic.	-0,569	-1,671	-0,971	-1,030	-0,732	-1,207	-0,969	-0,600
	(0,349)	(0,348)	(0,351)	(0,361)	(0,362)	(0,353)	(0,370)	(0,502)
services	-1,205	-1,277	-1,032	-1,022	-1,329	-1,270	-1,029	-0,531
	(0,030)	(0,030)	(0,042)	(0,047)	(0,045)	(0,045)	(0,042)	(0,052)
house non-durables	-2,154	-2,196	-2,197	-2,187	-2,475	-2,077	-2,070	-2,358
	(0,732)	(0,744)	(0,795)	(0,850)	(0,864)	(0,894)	(0,872)	(0,871)

Note: Standard errors are in parentheses

We have tried other possibilities such as to control for latent effects with estimations in levels by instrumenting expenditure with first differentiated income. These level estimations will provide inconsistent estimates as far as income is correlated with the latent effects and with infrequency.

We now present elasticity estimations for the rank three specification (see now columns 7 and 8 in Tables 2.1 and 2.2). From the above analysis, we detected the presence of measurement errors. However, we determined that they displayed an invariant in time nature and hence, first differences estimations canceled their effects. Under this assumption, Maximum Likelihood estimation applied to equations in first differences will provide consistent estimates (column 7). If not, Non-linear IV (minimum distance) procedure will not give consistent estimates. The repeated error measurement procedure of Hausman, Newey and Powell (1995) will provide consistent estimates (column 8). Adjusted expenditure for this estimation is constructed by regressing expenditure on the previously described socioeconomic variables and dummies of education and age and on future consumption.

ML estimation gives significance of the linear and quadratic expenditure terms except for transport and communications (see column 7 in Tables A.2.1 through A.2.8). For these results, we do not observe significant differences on expenditure and price elasticities between the rank 2 and rank 3 specifications for the first differences estimations. Recall that the presented parameters are obtained evaluating elasticities at the mean for an average consumer. However, dependence of expenditure elasticities on total expenditure determines a distribution upon these parameters that might be relevant for analysis on welfare when implementing tax reforms. Also, the error measurement procedure gives similar expenditure elasticities although most of the parameters from whom they are derived are non-significant. This result suggests that adjusted expenditure on the used exogenous household characteristics and future expenditure is not such a good proxy for current expenditure.[31]

2.5.2 - Theoretical restrictions

Once consistency of estimates is ensured, we focus our attention in obtaining a set of parameters that verify all the integrability conditions. In fact, these theoretical restrictions may be accomplished if we want to use the derived parameters for different simulation purposes that need to use utility or cost functions.

When estimating a complete demand system, the additivity condition with respect to expenditure is directly verified to keep full rank to N-1, being N the number of goods we are considering. The parameters of the last equation can be recovered from this property. Moreover, all the above estimations include directly the homogeneity property which in fact is regularly accepted on all empirical works. From the initial parameters obtained for the estimation in first differences, we apply symmetry by Minimum Chi-Square and reject

[31]The inclusion of lags of expenditure instead as regressors when constructing adjusted expenditure does not change the results significantly.

this integrability condition.[32] The set of expenditure and price elasticities derived from the parameters that verify symmetry are also shown in column 6 in Tables 2.1 and 2.2. Expenditure elasticities do not differ very much from those obtained before whereas price elasticities come up all as upwardly biased for the categories with a high percentage of zeros and downwardly biased for the rest except for services and non-durables.

The rank three specification carries an additional integrability condition which implies the same polynomial structure in *ln(X/P)* for all expenditure shares. We analyze if this is a strong assumption for our data. Comparing all linear and quadratic expenditure parameters, we observe that despite the differences among the parameters for each equation, their ratios are very close except for food and transport and communication, although the latter parameters are not significant (see appendix A.2). We have also imposed the linear integrability restriction $\varepsilon=\delta_i * \beta_i$ on all categories and derived ε by minimum chi-square. The χ^2 (7) test yields a relative high statistic (206.5) for ML estimates which must be attributed to the high significance of the linear and quadratic parameters. The ratios obtained from repeated measurement estimates are non-significant due to the poor precision of the estimates. Despite the rejection of the integrability restriction, the similarities between rank two and rank three specifications suggest that a rank two demand system is not a bad choice to describe demand analysis for our data since here we are not interested either on welfare analysis or in the distribution of expenditure elasticities.

---

[32]Imposing symmetry by the Minimum Chi-Square method, we express $\theta = K\beta$, being θ the non restricted parameters and β the symmetry restricted parameters. The symmetry parameter estimates can be obtained by minimizing the function:

$$argmin_{\beta}(\theta^* - K\beta)' \Sigma^{-1}_{\theta*} (\theta^* - K\beta).$$

where $\theta*$ is an estimate of the unrestricted parameter vector θ and $\Sigma^{-1}_{\theta*}$ is the inverse of an estimate of the variance-covariance matrix of $\theta*$.

The minimized value of this expression follows a Chi-squared distribution with m degrees of freedom. An estimate for the covariance matrix of $\beta*$ is $(K' \Sigma^{-1}_{\theta*} K)^{-1}$. We reject symmetry since we obtain a chi-square of 88,68 with 28 d.f.

2.5.3 - Elasticities from other studies. A comparison.

We finally present in Tables 2.3 and 2.4 a comparison between the derived expenditure and price elasticities from our selected best model together with parameters from other studies, also for the Spanish economy. The sets of parameters are derived from different functional forms (LES or AIM) and are directly comparable. Even though differences in magnitude among the studies suggest some changes in the behavior of the households, we must notice that an important part of the variation may be due to aggregation when constructing the expenditure categories as well as the sort of data used in each study and the econometric treatment of the model.

Nevertheless, a straightforward comparison reflects that necessities and luxuries are identified as such on all studies except rents and house keeping which we identify as necessity whereas the other studies as luxury. We must also notice the similarity in the results between our estimates and those reported in column 2. In fact, both are obtained using individual data derived from the same survey, except from the fact that our study covers a different period and some aggregates are not exactly defined. Nevertheless, there is an important difference between both estimates. Parameters in the second column do not take into account latent effects. As a very general pattern, we might say that our estimates are more extreme. That is, we characterize necessities with a lower expenditure elasticity and luxuries with a greater one. A similar structure seems to describe price elasticities. In particular, it is worth to mention that we obtain an acceptable price elasticity for house non-durables,[33] whereas Labeaga and López (1996) derive a more intuitive value for clothing. However, we observe closer results if instead, we compare income and price elasticities derived from our pooled estimation and those reported in column 2.

[33]Its high price elasticity value may be explained in terms of the heterogeneity of goods included in this category.

Table 2.3 - EXPENDITURE ELASTICITES

	1	2	3	4
Food and non alc. beverages	0,70	0,76	0,72	0,48
Alcoholic beverages	1,05	0,88	-	0,54
Clothing and foodwear	1,62	1,32	1,29	0,86
Rents and house keeping	0,71	-	1,84	1,16
Fuel for housing	0,36	0,86	-	-
Transport and communications	1,67	1,13	1,99	0,87
Services	1,23	-	-	1,18
House non-durables	0,96	1,49	-	-

Table 2.4 - OWN-PRICE ELASTICITIES

	1	2	3	4
Food and non alc. Beverages	-0,77	-0,87	-0,47	-
Alcoholic beverages	-1,66	-1,03	-	-
Clothing and foodwear	-0,11	-0,89	-0,68	-
Rents and house keeping	-1,78	-	-0,9	-
Fuel for housing	-0,99	-0,53	-0,97	-
Transport and communications	-1,03	-1,27	-0,87	-
Services	-1,02	-	-	-
House non-durables	-2,19	0,14	-	-

1) Estimates of our best selected rank 2 model.

2) Estimates derived from an AIM on the ECPF for the period 1985-1989 (Labeaga and López, 1996).

3) Estimates derived from a L.E.S. on cross-sectional data using the 1981 EPF (López, 1986).

4) Estimates derived from a L.E.S. on cohort data constructed from income centiles from the EPC, covering the period 1977-1981 (Abadía, 1984).

2.6 - SUMMARY AND CONCLUSIONS

In this chapter, we assess the importance of controlling individual effects, both observable and unobservable, on the estimation of price and income elasticities. Individual observable effects are described with demographic and socioeconomic characteristics whereas the latent effects refer to unobserved features of the family. Moreover, we use the distinction between observed expenditure and desired consumption in order to capture the errors that may be associated to infrequency purchases. Besides of this infrequency effects, we consider that observed expenditure may differ from desired expenditure due also to stochastic errors in expenditure variables.

We specify and estimate consistently both a rank two and rank three Almost Ideal Demand System, and from the obtained parameters, we derive price and income elasticities. We use a sample at the household level drawn from the ECPF (1985-1991) panel data survey for the Spanish economy. We follow each household throughout 6 quarters. Given such a short profile, we assume that heterogeneity displays an invariant in time pattern. Using panel data, we are able to control for the different components of the error structure, and also, we may describe the bias pattern derived from each source of error. Finally, the obtained set of consistent parameters from our best selected model are compared with previous similar studies on the Spanish economy.

To control for the different sources of errors leads to an specification in first differences and an estimation with IV, using lags of first differences of expenditure as instruments for expenditure. Different specifications and estimations allow to test for endogeneity of expenditure and income as well as the effects of unobservable heterogeneity and infrequency. First of all, we reject endogeneity of expenditure and do not reject endogeneity of income. Furthermore, infrequency depends on the probabilities of purchase which are usually modeled as dependent on household demographics. Since we observe that in our data these variables are time invariant, we check out whether infrequency displays an invariant in time behavior and conclude that effectively, once we control for invariant in time elements, the effects of infrequency of purchase vanish. Besides, demand analysis usually detects correlation between the latent individual effects and expenditure due to omitted variables. We assess that effectively there is such a correlation and confirm that pooled estimations lead to biased income and price elasticities.

The obtained results for the rank two consistent specifications confirm the intuition about whether goods are necessities or luxuries. Price estimates, although some of them are not relevant, are jointly significant and present the expected sign and size. These estimations that control for all the components of the whole error structure add some evidence on the direction of the bias derived when not controlling for any of the different sources of errors. Hence, we are able to describe that the latent effects tend to bias downwardly the income parameters and in the opposite direction price parameters. The expected effects of the error derived from the policy of purchase are in the opposite direction, especially on those categories with a higher incidence of zeros. Nonetheless, we observe that this pattern is only followed by housing, domestic fuel, services and house non-durables, and from this evidence, we conclude that a problem of errors in variables different than infrequency purchases is also present on our data, especially on the other categories.

Moving to rank three specifications, we observe that income and price estimates do not differ significantly from those obtained when restricting the rank of expenditure once we control for the presence of individual unobservable effects and infrequency. The high significance of income parameters seems to be the reason for a rejection on the rank three integrability condition. Therefore, we conclude that proxying demand analysis for the Spanish economy assuming a rank two for the demand system does not seem to be very restrictive using the ECPF and these categories of expenditure. However, dependence of the elasticities of the different categories on household total expenditure, determines the possibility to derive a distribution for expenditure parameters, which may be used on welfare analysis when simulating tax reforms.

REFERENCES

ABADIA, A. (1984): "Un sistema completo de demanda para la economía española". *Investigaciones Económicas*, 25, 5-18.

AMEMIYA, Y. (1985): "Instrumental variable estimator for the non-linear errors-in.variables model". *Journal of Econometrics*, 28, 273-90.

ARELLANO, M. and BOVER, O. (1990): "La econometría de datos de panel". *Investigaciones Económicas*, 14, 3-45.

ATKINSON, A.B., GOMULKA, J. and STERN, N.H. (1990): "Spending on alcohol: Evidence from the Family Expenditure Survey 1970-1983". *Economic Journal*,100, 808-827.

ATTFIELD (1985): "Homogeneity and endogeneity in systems of demand equations". *Journal of Econometrics*, 27, 197-210.

BALESTRA, P. and NERLOVE, M. (1966): "Pooling cross-section and time series data in the estimation of a dynamic model: the demand for natural gas". *Econometrica*, 34, 585-612.

BANKS, S., BLUNDELL, R. and LEWBEL, A. (1997): "Cuadratic engel curves, welfare measurement and consumer demand". *The Review of Economics and Statistics*, 79, 527-539.

BLUNDELL, R. and MEGHIR, C. (1987a): "Engel curve estimation with individual data". *The practice of econometrics*. Heijmans and Neudecker eds.

BLUNDELL, R. and MEGHIR, C. (1987b): "Bivariate alternatives to the Tobit model". *Journal of Econometrics*, 34, 179-200.

BLUNDELL, R., PASHARDES, P. and WEBER, G. (1993): "What do we learn about consumer demand patterns form micro data?". *American Economic Review*, 83, 570-97.

BROWNING, M. (1991): "A Simple, Non Additive Preference Structure for models of household behavior over time". *Journal of Political Economy*, 99, 607-637.

BROWNING, M and MEGHIR, C. 1991: "The effects of male and female labor supply on commodity demands". *Econometrica*, 59, pgs 925-951

CHAMBERLAIN, G. (1982): "Multivariate regression models for panel data". *Journal of Econometrics*, 18, 5-46.

CHAMBERLAIN, G. (1984): "Panel Data". In Z. Griliches and M.D. Intriligator eds. *Handbook of Econometrics*. Vol. II. Elsevier Science Publishers.

COLLADO, M.D. (1993): "Separability and aggregate shocks in the life-cycle model of consumption: Evidence from Spain". Preliminary draft. *CEMFI and LSE.*

CUTANDA, A. (1995): "Consumo y exceso de sensibilidad a la renta: Evidencia para el caso español a partir de un pseudo-panel de la ECPF". Working Paper 9516, *Universitat de Valencia.*

DEATON, A. and MUELLBAUER, J. (1980a). *Economics and Consumer Behaviour.* Cambridge University Press.

DEATON, A. and MUELLBAUER, J. (1980b): "An almost ideal demand system". *American Economic Review*, 70, 312-26.

GALLANT, A.R. (1981): "On the bias in flexible functional forms and an essentially unbiased form (The Fourier flexible form)". *Journal of Econometrics*, 15, 21-245.

GARCIA, J. and LABEAGA, J.M. (1996): "Alternative approaches to modeling zero expenditures: an application to Spanish demand for tobacco". *Oxford Bulletin of Economics and Statistics*, 58, 3, 489-506.

GORMAN, W.M. (1981): "Some Engel curves". In A.S. Deaton ed. *Essays in the theory and measurement of consumer behaviour.* Cambridge University Press.

HAUSMAN, J.A., NEWEY, W.K. and POWELL, J.L. (1988): "Consistent estimation of non-linear errors-in-variables models with few measurements". Mimeo. *MIT.*

HAUSMAN, J.A., NEWEY, W.K. and POWELL, J.L. (1995): "Non-linear errors in variables. Estimation of some Engel curves". *Journal of Econometrics*, 65, 205-33.

HAUSMAN, J.A. and TAYLOR, W.E. (1981): "Panel data and unobservable individual effects". *Econometrica*, 49, 1377-98.

HAUSMAN, J.A. and WISE, D. (1979): "Attrition bias in experimental and panel data: The Geary income maintenance experiment". *Econometrica*, 47, 455-473

HSIAO, C. (1986). *Analysis of panel data.* Cambridge University Press.

INE (1985): *Encuesta Continua de Presupuestos Familiares. Metodología.* Instituto Nacional de Estadística. Madrid.

KAY, J.M., KEEN, M. and MORRIS, N. (1984): "Estimating consumption from expenditure data". *Journal of Public Economics*, 23, 169-82.

KEEN, M. (1986): "Zero expenditures and the estimation of Engel Curves". *Journal of Applied Econometrics*, 1, 277-86.

LABEAGA, J.M. (1992): "A dynamic panel data model with limited dependent variables: an application to the demand for tobacco". Documento de Trabajo, 92/01, *UNED*.

LABEAGA, J.M. and LOPEZ, A. (1996): "Flexible demand system estimation and the Revenue and Welfare Effects of the 1995 Vat Reform on Spanish households". *Revista Española de Economía*, 13, 181-197.

LABEAGA, J.M. and LOPEZ, A. (1997): "A study of petrol consumption using Spanish panel data", *Applied Economics*, 29, 795-802.

LEE, L.F. and PITT, M.M. (1986): "Microeconomic demand systems with binding non negativity constrains: The dual approach". *Econometrica*, 54, 1237-42.

LIVIATAN, N. (1961): "Errors in variables and Engel curve analysis". *Econometrica*, 29, 336-362.

LOPEZ, A. (1998): "Unobserved heterogeneity and censoring in the demand for health care". *Health Economics*, forthcoming.

LOPEZ, E. (1986): "La estructura del consumo en España en 1981. Una aplicación del modelo lineal de gasto". *Cuadernos de Economía*, 14, 89-106.

LOPEZ-SALIDO, J.D. (1993): "Consumo y ciclo vital: resultados para España con datos de panel". *Investigaciones Económicas*, 17, 285-312.

LORENZO, M.J. (1987): "Sistemas completos de demanda para la economía española: una aplicación a la imposición óptima". *Investigaciones Económicas*, 12, 83-131.

MARSHALL, A. (1980). *Principles of Economics*. MacMillan eds.

MATAS, A. (1991): "La demanda de transporte urbano". *Investigaciones Económicas*, 15, 249-267.

MEGHIR, C. and ROBIN, J.M. (1991): "Frequency of purchase and the estimation of demand systems". *Journal of Econometrics*, 53, 53-85.

MOLTO, M.L., REIG, E. and URIEL, E. (1990): "La demanda de productos alimenticios en la Comunidad Valenciana: un análisis de corte transversal". *Investigaciones Económicas*, 14, 149-164.

MONES, M.A., SALAS, R. and VENTURA, E. (1992): "Consumption, real after tax interest rates and income innovations. A panel data analysis". Economics Working Paper, 18. *Universitat Pompeu Fabra.*

MUNDLAK, Y. (1978): "On the pooling of time series and cross-section data". *Econometrica*, 46, 69-86.

PASHARDES, P. (1993): "Bias in Estimating the Almost Ideal Demand System with the Stone Index approximation". *Economic Journal*, 103, 908-915.

PASHARDES, P. and BAKER, P. (1991): "Dynamic Specification and Household Characteristics in Demand Analysis". Working Paper, 1991-11, *Institute for Fiscal Studies.*

PHLIPS, L (1974). *Applied Consumption Analysis*. North Holland, Amsterdam and Oxford.

POLLAK, R.A. (1969): "Habit formation and dynamic demand functions". *Journal of Political Economy*, 78, 745-63.

POLLAK, R.A. and WALES, T. (1981): "Demographic variables in demand analysis". *Econometrica*, 49, 1533-1552

PUDNEY, S. (1989). *Modelling individual choice. The econometrics of corners, kinks, and holes*. Basil Blackwell. Oxford.

RAYMOND, J.L., OLIVER, J. and PUJOLAR, D. (1995): "El comportamiento del ahorro familiar a partir de la Encuesta de Presupuestos Familiares 1990-1991". *Papeles de Economía*, 65, 196-211.

SUMMERS, R. (1959): "A note on least squares bias in household expenditure analysis". *Econometrica*, 27, 121-134.

WHITE, H. (1980): "A heteroskedastiticy.-consistent covariance matrix estimator and a direct test for heteroskedasticity". *Econometrica*, 48, 817-838.