

Economics Working Paper 52

**The Demand for Private Transport in
Spain: A Microeconometric Approach**

Ángel López*

Universitat Pompeu Fabra

December 1993



UNIVERSITAT POMPEU FABRA

*Balmes, 132
Telephone (343) 484 97 00
Fax (343) 484 97 02
08008 Barcelona
e-mail econwp@upf.es*



Economics Working Paper 52

**The Demand for Private Transport in
Spain: A Microeconometric Approach**

Ángel López*

Universitat Pompeu Fabra

December 1993

Keywords: Panel data, Zero expenditures, Infrequency of purchase, Vehicle use, Petrol demand.

Journal of Economic Literature classification: C23, C25, C21, D12, H31.

*I am grateful to John Micklewright, José María Labeaga and the researchers at CEMFI for useful comments to previous versions of this paper. All errors and omissions are mine.

Universitat Pompeu Fabra, Balmes 132, 08008 Barcelona. Tel +43 3 5422708. Fax +34 3 4849746. E-mail angel@upf.es

Abstract

Repeated observations on the expenditure habits of households permit an attractive treatment for the zero expenditures in general and infrequent purchases in particular that contaminate the information in a typical cross-section budget survey. This paper exploits in this sense the panel features of the Spanish Encuesta Continua de Presupuestos Familiares in the context of vehicle ownership and associated rate of usage in order to contribute with a microeconomic study of the consumption of private transport to the recent trend in empirical demand studies at the household level for the Spanish economy.

1 Introduction

Together with alcohol and tobacco, motor fuel bears both excise duties and VAT in most OECD countries. Likewise, the purchase of new vehicles is often subject to luxury VAT rates or other purchase taxes. Thus private road transport has traditionally been one of the consumption categories yielding an important share of revenue to Exchequers. Efforts in the direction of studying the characteristics of demand for private transport are fully justified from this angle but, in addition, the debate¹ over the use of taxes as a means of regulating environmental externalities has increased the policy interest of academic research into this area. For instance, apart from the changes that the imposition of minimum excise duties over the EC has brought upon petrol prices, the introduction of an energy tax in the EC has increased the price of a barrel of oil 3 dollars in 1993 and will increase it by one dollar a year until the year 2000. As far as the particular case of Spain goes, the de-regulation of the supply of petrol at road outlets and the recent devaluations² of the peseta increase the probability of further changes in the price of motor fuel, as recent events have confirmed.

The research in this paper aspires to providing some useful evidence for the analysis of the consequences of those changes

¹ See Pearson and Smith (1991).

² Spain imports its crude and pays hard currencies for it.

on households' demands and welfare. This purpose generally requires the estimation of a demand equation in order to calculate the elasticities of interest and the computation of welfare measures. In this sense, this research contributes to the series of applied work on demand using Spanish household data³.

The nature of the demand for private transport differs from that of other non durable goods in that it is a joint demand for a vehicle and its carburant. Demands of this type are best viewed within a household production framework in which the household derives utility from the consumption of private transport requiring vehicle and petrol as previous inputs. Thus, unlike models trying to describe the relationship between aggregate consumption, prices and aggregate income, micro level exercises have to address the household choice over durable stocks. Since the latter are best viewed as investment, rather than consumption decisions, provision of an adequate modelling framework can hardly be done in the same terms as those used for short-lived goods⁴. Thus, although there are investigations in which type of durable tenure and rate of usage are addressed jointly⁵, two types of studies usually emerge in the literature: those concentrating on ownership and those addressing the rate of usage conditional on the existing stock of durable goods⁶.

Two reasons why the demand for petrol should ideally be treated in conjunction with the demand for vehicles are, firstly, that

³ See for instance Labeaga and López (1992, 1993) for indirect tax simulations on systems of demand equations.

⁴ The demand for a vehicle presumably depends much more on expectations about future circumstances (changes in place of residence, job prospects, future prices of petrol etc.) than that for petrol conditional on having a vehicle.

⁵ Dubin and McFadden (1984) model type of heating system and subsequent demand for energy for the U.S. King (1980) studies the demand for housing services associated to three housing tenure regimes simultaneously.

⁶ Examples for the U.K. are Baker et al. (1988) for the former and Baker et al. (1989) for the latter.

the rate of usage and petrol requirement will be intimately associated with the type of vehicle in the household and, secondly, that by doing so we would be provided with an estimate of the effect of prices on the durable stock rather than the sole short run response which can be obtained otherwise. Unfortunately, the data available for this study do not offer direct information on ownership, let alone type of vehicle. As I shall explain, this requires a rather indirect approach in order to obtain information about the distribution of petrol expenditures conditional on using a vehicle and, in these circumstances, a joint approach seems impractical even if an adequate theoretical model was readily available.

Notwithstanding these difficulties, this research goes some way towards the aim stated before by studying the determinants of ownership and by estimating a single equation model of the demand for petrol conditional on vehicle use. In Section 2 I comment the data from the Encuesta Continua de Presupuestos Familiares (ECPF) (1985-89) in terms of both problems of measurement and features of private transport consumption. These features are analysed in more detail in section 3 through the estimation of econometric models of the probability of ownership and the rate of usage conditional on ownership. Section 4 gives concluding remarks and discusses the policy implications of this research.

2 Problems of measurement and main features of transport consumption in the ECPF (1985-89)

The ECPF is a quarterly budget survey of Spanish households. Its sample consists of 3200 households which are rotated in a 12.5% every quarter thus allowing the possibility of constructing panels of household for periods of up to eight quarters⁷. The information on consumption of private transport basically

⁷ A detailed analysis of the characteristics of the ECPF (1985-89) in terms of quality of response and usefulness for demand analysis can be found in chapter 3 of López (1993).

consists in the record of petrol purchases in the week of the interview. Although the data available for this study also contains information on vehicle purchases, this information cannot unambiguously indicate the existence of a carbureted vehicle, since it comprises non-carbureted vehicles⁸.

In the cross-section sample associated to any quarter, households who do not own a vehicle will be indistinguishable from households who regularly consume petrol but happen not to have purchased it during the monitoring period (the period over which the interviewer asks the household to report purchases, which is a week for petrol expenditures): their record on petrol expenditure will be zero. The microeconometrics literature would also consider a third cause for these zero records, that of corner solutions i.e. the household chooses a level of zero consumption as a result of his/her utility maximisation problem. In the context of petrol demand, however, it is unreasonable to abstract from the fact that such decisions are conditional on vehicle availability. Thus, once the costs of purchasing a vehicle (plus insurance etc.) have been incurred, it does not make much sense to allow for the possibility of consumers being priced out of the consumption of petrol. As an identifying assumption, therefore, I shall maintain that zeroes are due either to not having access to a vehicle or, for those who have it, purchases being made with a frequency longer than a week⁹. The crucial issue for the analysis of this data set is therefore how households with a vehicle may be told apart from those without and how can a measure of consumption (as opposite to purchases) be obtained.

⁸ For a detailed description of these issues see chapter 6 of López (1993).

⁹ The validity of this assumption can, in principle, be tested since the three causes for zeroes can be nested within a likelihood function (see Blundell and Meghir 1987 and Pudney 1989). In any case, there are studies in which similar a priori assumptions have been made (see Meghir and Robin 1992).

2.1 Imputing use/ownership of a vehicle

Clearly, a positive expenditure on petrol in any quarter over the number of quarters that a particular household collaborates in the survey should indicate their use of a vehicle. As the margin for infrequent petrol purchases to hide the existence of a vehicle decreases with the number of quarters a household is interviewed (i.e. if the household uses a vehicle, the more times it gives information on expenditures, the more likely petrol expenditures are to show up), it would be ideal to base inferences about ownership on the group of households that stay in the survey during the stipulated eight quarters. However, the problem of attrition (i.e. lack of completion of the sequence of eight interviews) in the ECPF 1985-89 renders such group very small (1505 households)¹⁰ so a compromise between sample size and purge of infrequency looks rewarding. The next piece of evidence would suggest that four quarters may be the threshold above which the consumption of petrol (and therefore the use of a vehicle) is revealed through purchases in at least one quarter.

TABLE 1 BREAKDOWN OF POSITIVE EXPENDITURE RECORDS FOR PETROL IN AT LEAST ONE QUARTER DURING COLLABORATION BY NUMBER OF QUARTERS COLLABORATED

No. quarters	1	2	3	4	5	6	7	8
% positive in at least one quarter	42	55	62	64.5	64.1	65	68	72

N.B. Table obtained from 13711 households interviewed from one to eight times over the 20 quarters in the period 1985-89. No household is in two groups at the same time.

The figures in the table above show the number of households who record a positive purchase on petrol in at least one quarter as a percentage of the total of households completing each sequence of interviews. For instance, 62% of households who collaborate three times with the survey would thus seem to have purchased petrol (and used a vehicle).

¹⁰ See chapter 3 of López (1993).

If the process determining attrition is not related in any way to the determinants of vehicle ownership, it may be reasonably expected that the percentage of households using a vehicle (and therefore consuming petrol) to be constant across each of the eight groups above with differences in the observed percentages due only to infrequency of purchases. Given this, Table 1 suggests that the rate of purchases of petrol in at least one quarter ceases to increase substantially as we pass the threshold of households collaborating four quarters¹¹ and is centered at around 2/3 of the total of households in each of these groups¹². Consequently, in order to minimise the scope for infrequent purchases to induce erroneous inferences while keeping a sufficient number of households, I shall base the subsequent analysis on households which collaborate during at least four quarters. This solves the problem of identifying motor-vehicle users, i.e. these are the households that show a positive petrol record at any time during participation in the panel¹³.

¹¹ The difference of 4% between households collaborating seven and eight times raised suspicion as to whether attrition is random (or at least uncorrelated with the determinants of vehicle use). An examination of the observable characteristics of these two groups did not offer any alternative explanation.

¹² This figure fares reasonably well against the information in the OECD survey of Spain for 1992 which publishes a figure of 263 private vehicles for every thousand inhabitants. This would yield a rate of 66% of households owning a car if the average household size was of 2.53 persons and this is one person less than our sample average household size. However, we must recall that the ECPF is a household survey where institutional consumers are not represented and where small size households are underrepresented. This may account for the divergence between these two figures.

¹³ This is not the unique way to impute vehicle use/ownership in a data set with these characteristics. An alternative would have been to use data from the Spanish Encuesta de Presupuestos Familiares (EPF) which contains information on the existence of a vehicle and approach the probability of ownership as a function of demographic and other variables common to the ECPF. This could have then been used to impute ownership to the households in the ECPF (see Arellano and Meghir 1988). The EPF data available at the time this study was carried out is for the period 1980-81 and since the macroeconomic circumstances for the latter part of the decade were radically different, a combination of the two approaches may still have been necessary.

Having obtained an indicator of ownership, I now turn to two related questions: How important is the problem of infrequent purchases and how big are the chances of mistaking for infrequent purchases the fact that the household did not own a vehicle until some point in time during his collaboration spell? These two issues are even more relevant when we take into account that the eighties were a period of huge growth for the Spanish economy and that the number of vehicles in circulation went up from 11,716,000 in 1985 to 14,870,000 in 1989¹⁴.

The information offered by the records of expenditure on new or second hand vehicles is of use now because any vehicle purchase during the surveyed period will necessarily be reflected in the expenditure records (the monitoring period for these expenditures coincides with the survey period i.e. one quarter). This permits the detection of situations in which the household does not buy petrol until the time when it buys a vehicle from genuine infrequent purchases. The following table shows the evidence on such contingencies by providing a detailed account of the number of times a positive expenditure on petrol is recorded for households who use a vehicle and what is the number of cases in which no infrequency of purchase in petrol is observed after a positive record for "vehicles".

¹⁴ Source: Anuario del Instituto Nacional de Estadística. Madrid. The figures include motorcycles and lorries.

TABLE 2 NUMBER OF QUARTERS WITH OBSERVED POSITIVE PURCHASES AND NUMBER OF POTENTIAL FALSE CASES OF INFREQUENCY WITH CORRESPONDING PERCENTAGES BY PERIODS OF COLLABORATION

	Number of collaborated quarters				
	4	5	6	7	8
Number of positive petrol records in at least one quarter	1301	1205	992	632	1077
In only 1 quarter	195	142	81	55	104
In only 2 quarters	233	145	106	64	77
In only 3 quarters	349	170	122	74	94
In only 4 quarters	524	293	159	63	101
In only 5 quarters		455	221	100	109
In only 6 quarters			303	118	168
In only 7 quarters				158	206
In all 8 quarters					218
Cases of "False infrequency"	112	95	90	43	76
Proportion observed infrequency	.60	.62	.69	.75	.79
Proportion adjusted infrequency	.51	.54	.60	.68	.72

1) "False infrequency" is the term I use for households who not having recorded positive expenditures on petrol do so at every period starting from the period in which a record on "vehicles" is observed.

2) Proportion of observed infrequency and adjusted infrequency refer to the percentages of motorised households for which at least one zero record is observed and such percentage adjusted for the potential cases of "false infrequency".

3) Table obtained from households who use a vehicle on the criterion discussed previously. No household belongs to two groups at a time.

The first row of figures in the table show the number of households to which ownership is imputed for each number of interviews, since they have one positive record on petrol in at least one of these interviews. The subsequent rows show how many of these households record positive petrol expenditures in at most one quarter, two quarters and so on. The figures in the "false infrequency" row correspond to the number of households who have a consecutive spell of zeroes, a positive record on "vehicles" and then a consecutive spell of positive records up to the end of their collaboration, which introduces the possibility that they did not use a vehicle up to the point in which the first positive record is detected. The two rows headed proportion of observed and adjusted infrequency show the proportion of households recording at least one zero including and excluding the cases of "false infrequency" respectively. Thus 60% of households who collaborate four times, for instance, have

recorded one zero expenditure which is caused either by infrequent purchases or by not using a car, but at least 85% (51% of 60%) of these cases are caused exclusively by infrequency (under my assumption that there are no corner solutions), since an alternation of zero records and positive ones is observed.

The figures above seem to suggest that although infrequent purchases are made by the absolute majority of households, the cases in which a positive record appears in all the interviews is always the most numerous group. This may mean that the week is not so much an inadequate reference period for petrol for a large part of the population but it also reveals how there exists a high degree of heterogeneity in the underlying purchasing patterns. Further, the number of cases in which there exists the possibility of mistaking infrequent purchases for changes of "ownership" state seem not to be large in any of the groups of households above but, in any case, measurements on the continuous variable petrol expenditure should be taken at the end of the collaborating spell for every household to whom I impute the use of a vehicle in order to make sure that zeroes are genuine cases of infrequency of purchase¹⁵.

2.2 The measurement of the rate of usage conditional on ownership

The problem posed by infrequent purchases is not as easily overcome when looking at the levels of expenditure as when studying ownership. The fact is that observed expenditures are "contaminated" by the purchase strategy of each household and, typically, observed purchases do not correspond to underlying desired consumption unless the household purchases petrol at every reference period (i.e. every week in this case). This is noted by Pudney (1989 pp. 173-174):

... a good which is purchased regularly once every

¹⁵ The possibility of moving from having a vehicle to not having a vehicle is neglected on the grounds of casual observation.

four weeks will have a 75 per cent probability of no purchase being observed in a one-week enquiry, and a 25 percent probability of purchase being observed in a quantity four times as large as the underlying average weekly rate of consumption.

Therefore positive expenditures overstate desired consumption and, obviously, zero expenditures do the opposite.

It is possible to define expected observed expenditure on a good j for a household h , y_{hj} , in the following way

$$E(y_{hj}) = E(y_{hj}/D_{hj} > 0) P_{hj} + E(y_{hj}/D_{hj} < 0) (1 - P_{hj}) ; \quad (1)$$

where D_{hj} is the latent binary variable that determines whether purchases are made or not during the reference period (i.e. $D=0$ for no purchases and $D=1$ for purchases) and P_{hj} is the probability of D being one. Further, it is generally assumed that observed expenditures and desired consumption, y_{hj}^* , are equal in expectation, $E(y_{hj}^*) = E(y_{hj})$, that is, the average rate of expenditure over a sequence of observations for a household coincides with its average rate of consumption.

In these circumstances, the link between the two counterparts may be written in the following way (see Blundell and Meghir 1987 and Pudney 1989):

$$y_{hj} = \frac{y_{hj}^*}{P_{hj}} + v_{hj} ; \quad (2)$$

As P_{ij} can be considered an individual depreciation rate¹⁶, the expression above suggests that observed expenditures be deflated by the probability of observing a purchase to obtain a measure of desired consumption. With cross section data such probability is not directly observable for every household and therefore it has to be estimated¹⁷. But the ECPF panel has two advantages in this sense. First, it provides us with a consistent estimator of the probability of purchase conditional on the household being an user/owner simply by looking at the number of times that an user/owner household is recorded spending on petrol out of the total number of collaborated periods¹⁸. Second, where households positively known to own a vehicle record zero expenditures at any single quarter, the absence of information on desired consumption can be overcome by looking at previous or subsequent positive records¹⁹.

These two advantages permit the construction of a measure of desired consumption for the sample of vehicle-user households: I take households at their last collaboration and deflate their observed petrol expenditures by the estimate of their probability of purchase. This deflator is, as mentioned before, the estimate

¹⁶ Blundell and Meghir (1987) point out that the assumption regarding the equivalence of expected desired consumption and expected observed expenditure may be unrealistic for individuals at the end of their life cycle and therefore provide a corresponding test. However, the commodity they study is clothing, which is much more long-lived than petrol, and therefore more likely to violate the assumption if the household is old.

¹⁷ See Deaton and Irish (1984), Pudney (1989) or Meghir and Robin (1992).

¹⁸ It is useful to point out here that the ECPF interviews are carried out during weeks chosen randomly every quarter.

¹⁹ Of course, this latter possibility is granted by the implicit assumption that one single expenditure record at any quarter is sufficient, when the contamination induced by infrequent purchases is removed, to reflect the underlying, constant, rate of consumption for every household i.e. there are no dynamics in the rate of consumption nor in observed expenditures.

of the probability of purchase in any quarter conditional on having observed a purchase i.e. on the household being an user²⁰. Where the observed sequence of vehicle-petrol purchases is zero up to a positive record on vehicles (i.e. a case of "false infrequency"), the count allows for the possibility of such case to represent a passage from non-owner to owner. Also, if the last record on petrol of the household is zero, I can replace the observation by the most recent previous one in which a positive expenditure was observed and deflate it in the same way. This procedure provides me with a cross-section (7887 households) in which both the problem of vehicle use and that of infrequent purchases are by-passed, at least at the time of providing descriptive evidence on ownership and rate of usage.

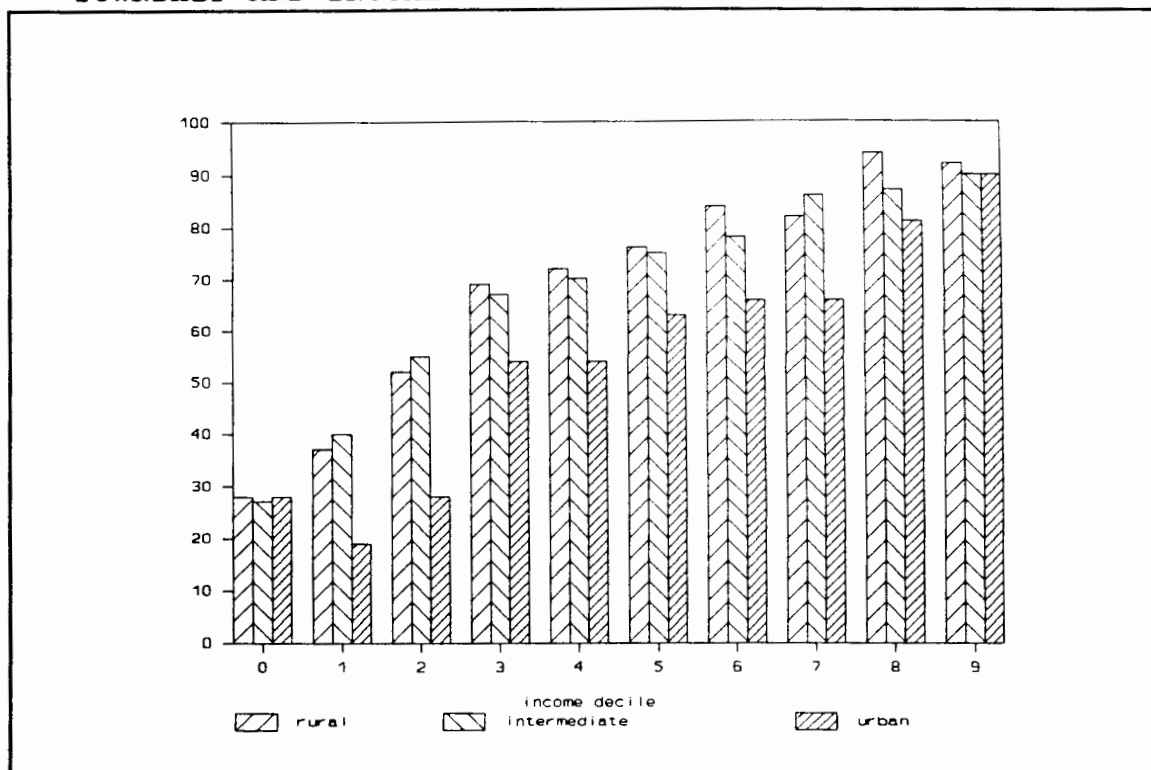
2.3 Some salient features of private transport consumption in Spain from the ECPF (1985-89) data

A study by Johnson, Mckay and Smith (1990) shows how density of population, region and income determined distinct ownership rates in the U.K. using data F.E.S. of 1986. These same variables seem to affect ownership also in Spain²¹, where the patterns in figure 1 can be observed from the data.

²⁰ Note that this probability is not the same as the probability of observing at least one purchase, i.e. the one modelled by the Probit model of section 3.

²¹ The data used are the 7887 households which collaborate from 4 to 8 times taken at their last collaboration. A breakdown of ownership rates by region is not possible since such variable is not available in the data.

FIGURE 1 PERCENTAGES OF VEHICLE OWNERS BY SIZE OF TOWNSHIP AND INCOME DECILE



N.B. Rural refers to townships of less than 10,000 inhabitants and city to townships of more than 50,000. Income is not adjusted by household size neither by other equivalence scale.

Clearly, there is a strong correlation²² between the level of income and the rates of ownership. This is hardly surprising as the costs of purchasing and running a vehicle, be it a motorcycle or a car, are by no means trivial. Other less intuitive but equally marked features are the consistently higher rate of ownership for rural households, only surpassed by households in intermediately populated areas for the second, third and seventh deciles of the income distribution. This may be a result of differences in availability of public transport between urban and rural areas. It is also conceivable that to use a vehicle in a big city is more "hazardous" than to do so in less densely populated areas, the reasons being traffic jams and difficulty in finding parking places. If urban households can do without

²² As it stands in the above graph, the relationship seems to be non-linear. In fact, it presents the shape of a quadratic in income with a negative coefficient on the squared term. This feature is also reflected in the probit results of next section.

public transport, they certainly have an incentive to do so in this sense. The fact that at the very top of the income distributions the rates of ownership are not so different seems to support this view, for very rich households are more able to afford skipping these two externalities by purchasing parking space on urban soil and, perhaps, devoting the use of the vehicle to trips outside city boundaries or for leisure purposes.

Concerning the rate of usage, the resulting distribution of "desired" petrol expenditures and that of positive observed petrol expenditures without adjustment are shown in histograms 1 and 2 respectively in the appendix. Both distributions seem to be bi-modal (modes around 11,000 and 26,000 pesetas) and present a long tail of outliers to the right. One possible reason for the bi-modal nature could be unobserved differences in the nature of the vehicle owned, notably whether it is a motorcycle or a car and within the latter whether it is a diesel or a high consumption engine, with motorcycle and low consumption car owners concentrating at the lower end of the distribution.

The next step is to measure petrol expenditure as a share of total expenditure²³. The share distribution thus obtained is depicted in histogram 3, where the tailing-off towards the right is still present but the bi-modal (the only mode now is around 3% of total expenditure) feature seems to have disappeared. If the above conjecture about the reasons for there being two modes is correct²⁴, this seems to suggest that the heterogeneity due to type of vehicle can be absorbed by differences in the level

²³ Total expenditure here refers to all expenditure on both durable and non durable goods.

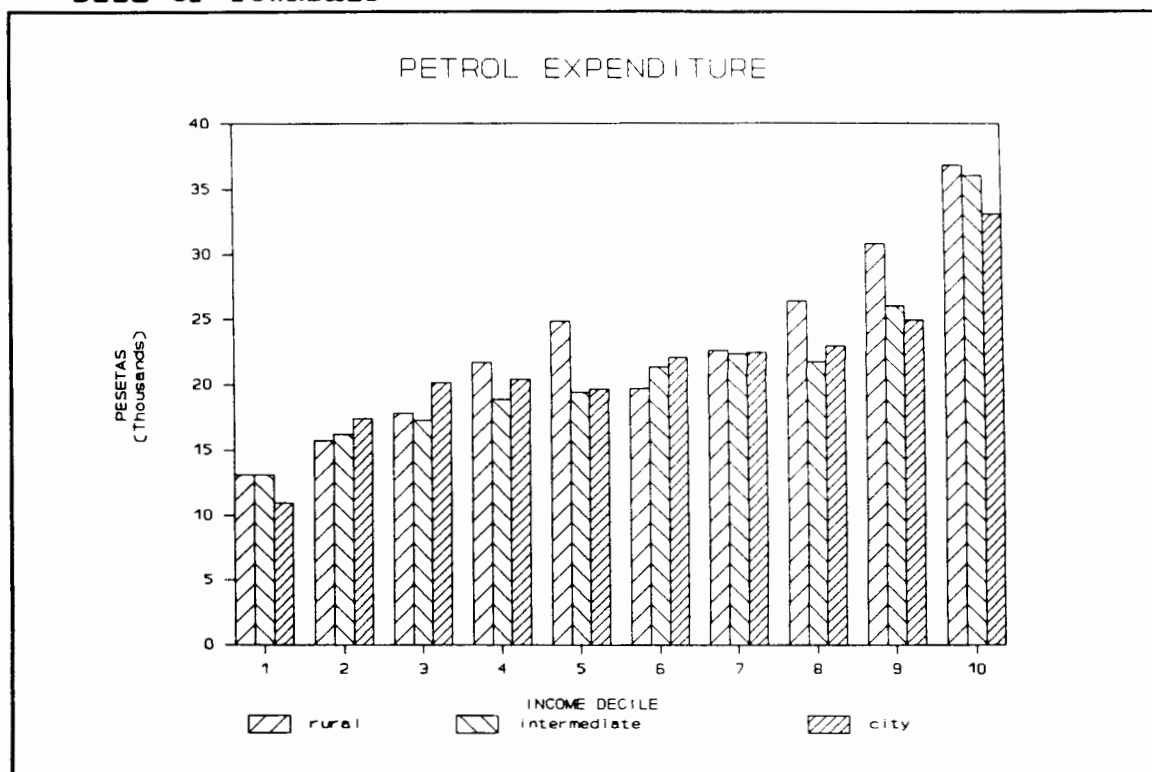
²⁴ It may be argued that the bi-modal feature may be due to having pooled together households surveyed at different periods in time so that differences in petrol prices would show up as several concentrations of mass in the distribution of expenditures. However, if the bi-modality was originated by heterogeneity in prices rather than quantities, it should still be present in the shares distribution because the price of petrol has diverged radically from the general pattern of inflation (this latter affects the denominator of the budget share).

of total expenditure to some extent. In order to look in detail at the share distribution I show the histograms of shares below and above 10% of total expenditure. Those below that benchmark, which make up nearly 90% of the whole share distribution, present a familiar bell-shape, and those above present the expected tailing-off. I have looked at the observable characteristics of these two groups in the hope of finding significant differences in terms of, perhaps, occupation or location but the results did not show any such difference. Where the occupation of the head of household has to do with transport or requires constant traveling (i.e. lorry drivers, delivery agents, salesmen or even shop-owners who travel to wholesalers²⁵) I would expect the share to be high but the set of explanatory variables in the data does not contain such information. With respect to the lower end of the shares distribution, I have found some extremely low expenditures that cannot possibly be destined to fuel a vehicle. As there is nothing particularly surprising about this since households may purchase petrol in order to light chimney fires and as I want to avoid the undue influence of outliers, I have dropped from the sample the bottom and top 1% of the shares distribution before proceeding with the analysis.

Like ownership, petrol expenditures rise with the level of income, and for a large part of the income distribution, rural households spend more on petrol, as shown in figure 2.

²⁵ I would expect the majority these of occupations to be included in the self-employment category but the way information is collected does not prevent employees to include petrol in the diaries even if it is an input that the firm reimburses subsequently.

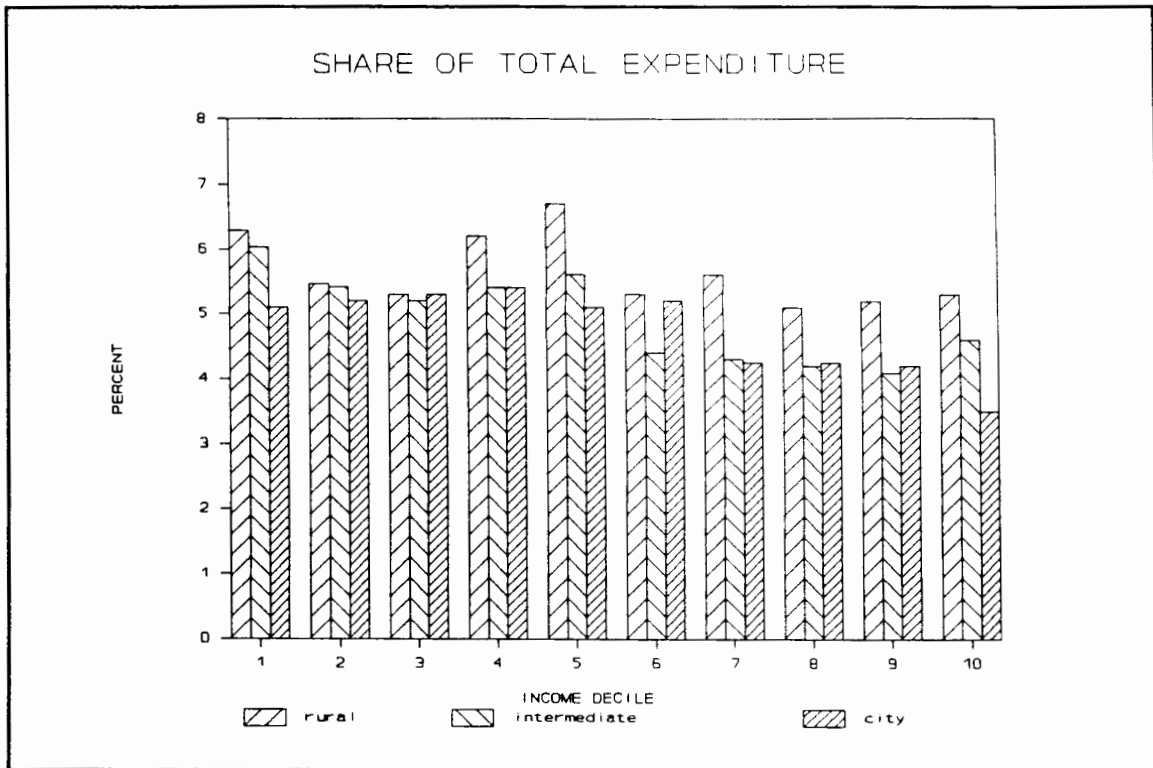
FIGURE 2 EXPENDITURE ON PETROL BY INCOME DECILE AND SIZE OF TOWNSHIP



N.B. Rural refers to townships of less than 10,000 inhabitants and city to townships of more than 50,000.

Whereas petrol expenditures in figure 2 are lower over some ranges of the income distribution for rural households, the budget share for the latter is never lower than for the rest of households, as figure 3 shows. This would suggest that private transport is more of a necessity for rural households than for the rest of households.

FIGURE 3 SHARE OF PETROL IN TOTAL EXPENDITURE BY INCOME DECILE AND SIZE OF TOWNSHIP



N.B. Rural refers to townships of less than 10,000 inhabitants and city to townships of more than 50,000.

Moreover, the general downward slope of the plot of shares against income deciles suggests that petrol is an overall necessity (among all vehicle users).

Thus, in general terms, the analysis of the data just presented would suggest that both ownership and rate of usage are strongly correlated with the level of income but the budget share of petrol falls with income. Size of township seems to have a positive impact on these two. In the next section I shall analyse ownership and rate of usage as functions of all the observable characteristics in the data set.

3 An econometric model for the probability of "ownership"

The last section showed how rich and rural households tend to have higher rates of ownership than the rest of households.

Clearly, there are more observables (and presumably a high number of unobservables too) affecting ownership (or vehicle use). A good way to summarise the effect of these (observables) is through the estimation of a probability of use model as a function of observable variables. Here, I shall use the PROBIT model. This model assumes that the ownership regime for every household is ruled by an underlying unobservable variable linear in parameters and with a normal additive disturbance of mean zero and variance one.

$$w_h = \alpha' z_h + v_h; \tag{3}$$

which yields an observable binary indicator

$$d_h = \begin{cases} 1 & \text{if } w_h > 0 \\ 0 & \text{otherwise;} \end{cases} \tag{4}$$

hence the sample likelihood is

$$L = \prod_{d_h=1} P(v_h > -\alpha' z_h) \prod_{d_h=0} [1 - P(v_h > -\alpha' z_h)]; \tag{5}$$

The following table presents the parameter estimates obtained by maximising the above likelihood function on the sample of 7730 households taken at their last collaboration. The binary variable is constructed according to the criterion of ownership discussed previously, that is, it takes the value of one if there has been a positive record for petrol during any of the panel waves and zero otherwise.

TABLE 3 PARAMETERS AND STANDARD ERRORS FOR PROBIT MODEL ESTIMATED BY MAXIMUM LIKELIHOOD

Variable		Variable	
Constant	-21.35 (4.54)	D u m m i e s	
Demographics		(continued)	
# Members	.289 (.036)	Children aged < 6	-.002* (.0014)
# Earners	.075 (.027)	6 < " < 14	-.0024* (.0012)
Dummies		14 < " < 23	-.0023 (.001)
Age < 35	.444 (.068)	Other dependents.	-.352 (.08)
35 < age < 50	.225 (.055)	Rural area.	.099 (.041)
age > 65	-.287 (.059)	City area.	-.385 (.050)
Unemployed	-.279 (.074)	Home renter	-.121 (.041)
Own employed	-.074* (.056)	Second home owner	.207 (.065)
Unskilled	-.285 (.055)	Expenditure	
White collar	.105* (.125)	Log total real income	2.67 (.720)
Not active	-.567 (.062)	" squared	-.078 (.025)
One parent	.004* (.249)	Log-Lkhood.	-3468
Pensioner	-.609 (.144)	Restr. Log-Lkhood.	-4167
Illiterate	.009* (.05)	Pseudo R ²	.167
Universitary	-.161* (.091)		

N.B. Parameters marked with an asterisk are not significant at the 5% level.

The degree of significance of parameters for the Probit seem to be high. The null hypothesis of joint insignificance is well rejected²⁶, the likelihood function is increased by 16% by including 24 additional parameters out of which 17 are individually significant:

3.1 Effect of demographic characteristics²⁷

Family size has a large effect on the estimated probability of use of a vehicle. With respect to the average family size, one additional member increases the predicted probability by some 10%. This may reflect two phenomena: one is that the bigger the household the bigger chances are that one of its members is a driver; secondly, private transport induces economies of scale with the number of passengers so that it pays to own a vehicle if the family is large.

As far as labour market status dummies are concerned, I did not expect that being unemployment or being out of the labour market (not active or retirement dummy) would have an unambiguous effect on the probability of using a vehicle. The reasons are that although the need to use private transport to travel to work (or to work itself) is not present for these people, the increase in free hours may lead to a use of a vehicle for leisure purposes and, if unemployed, for job hunting. From the results, however, it appears that those are effects one should expect to exist conditionally on there being an available vehicle. The negative sign on the estimated parameters for the unemployment, "not active", and pensioner dummies with their corresponding drops in

²⁶ The likelihood ratio test for the hypothesis of joint insignificance is equal to 1398 and it is distributed as a chi-squared (24).

²⁷ N.B. the effect of explanatory variables on the predicted probabilities for this model is given by the following expression

$$\frac{\partial \Phi(\alpha'z_h)}{\partial z_k} = \phi(\alpha'z_h) \alpha_k . \text{ This effect therefore depends on the}$$

configuration of the rest of explanatory variables.

predicted probabilities with respect to the base situation in table 6.4 suggest that these stata are associated to the absence of a private transport medium.

The inclusion of type of employment dummies responds to the idea that jobs requiring the use of a vehicle should be associated to one of the three categories (self-employment, white collar or unskilled). The lack of significance of all but one of these three estimated parameters does not throw much light on such a presumption except for the suggestion that unskilled workers are less likely to use private transport than the rest of workers (so they seem not to be employed in the transport business, at least in jobs involving purchases of petrol out of their own pocket).

The single-parent dummy is included in order to pick up differences in need for private transport with respect to a couple with children. If the single parent works, a vehicle may be necessary to take children to kindergartens (unless a baby-sitter is available) because the possibility of the other member looking after them is not present. In close relation, the existence of children (of any age) does not seem to be associated to the existence of a vehicle either.

Given the rest of conditioning variables, education dummies should principally pick up the differences in difficulty at the time of passing a driving test (in Spain it involves a written exam) but such effect seems not to be present in the data as the estimated parameters are not significant.

The importance of including population density dummies has been highlighted in the analysis of preceding sections. The results support my previous findings and predict a probability of vehicle use 12% higher for rural households with respect to urban ones *ceteris paribus*. The factors determining this difference should be found in the availability of public transport and the hazardous nature of urban motoring alluded to before.

There are reasons to believe that housing tenure status should be associated to vehicle use. Owning a house increases the level of lifetime wealth of a household. If income is correlated with lifetime wealth then, from the analysis in previous sections, I would expect vehicle ownership (not just use) to follow suit. Therefore I expected home renters to be less likely to own a vehicle than home owners. The results show evidence in this direction as home owners have a predicted probability some 2% higher than renters. If in addition the household owns a second house²⁸ then the predicted probability is some 6% higher.

3.2 Effect of real income

The level of real income has a non-linear positive (at a decreasing rate so the impact on the predicted probability is larger at small levels of income) effect on the probability of use. Table 6.4 shows how moving one standard deviation away to either side of the mean value of the logarithm of real income has a substantial effect on the predicted probability of use. In particular, the elasticity of the predicted probability to changes in (the logarithm of) real income at the mean is near the value of eight. This is not surprising since the use of a car (especially for user-owners) requires a substantial "sunk" payment plus a continued flow of running expenses such as insurance, repairs and, of course, carburant. These costs may act as a deterrent for low income households. Secondly, rich households will tend to consume more leisure and, to the extent that private transport is a complement to leisure, we may expect them to use/own a vehicle.

The following table charts the effect of explanatory variables on the predicted probabilities.

²⁸ I also expected a priori that if the second house is used for holiday purposes then a vehicle would be likely to be used to travel to and from it.

TABLE 4 EFFECT OF DEMOGRAPHICS AND REAL EXPENDITURE ON PREDICTED PROBABILITIES FROM THE PROBIT MODEL

Demographic	At mean expenditure	- 1 s.d. mean exp.	+ 1 s.d. mean exp.
Base	.86	.72	.94
Age < 35	.93	.85	.97
Age > 65	.78	.62	.90
Unemployed	.78	.62	.90
Not active	.69	.51	.84
Pensioner	.68	.50	.83
One parent	.86	.73	.94
Rural area	.88	.76	.95
City area	.76	.58	.90

N.B. The second and third columns are the effects of demographics at minus one standard deviation from the mean and plus one standard deviation from the mean of the logarithm of real expenditure distribution.

The characteristics of the base household are: aged between 35-50, blue collar employment with intermediate education, without children and home owner in a low density urban area.

3.3 Forecasting performance

Let us turn now to the forecasting performance of the model. Table 5 shows the pattern of correct and wrong predicted outcomes

TABLE 5 ACTUAL AND PROBIT PREDICTED OUTCOMES

	OWNERSHIP	NO OWNERSHIP
ACTUAL	5107	2623
CORRECT	4596	1547
WRONG	511	1076

N.B. I take as correct predictions those for which

$$\hat{P}(d_h=1) > .5 \text{ if } d_h=1, \hat{P}(d_h=1) < .5 \text{ if } d_h=0;$$

Within the group of owners, 90% of cases are predicted correctly. However, this percentage goes down to 59% for the group of non-owners. The overall rate of correct predictions is therefore nearly 80%. It is well known that when the percentage of zeroes and ones in the sample is not symmetric a probability threshold of 50% yields more correct forecast on the side of the more numerous group in the data. This is simply because the average predicted probability coincides with the proportion of ones in

the sample (the mean of the dependent variable). Thus if such mean is very high then it is hard to find a configuration of regressors that produces a predicted probability of less than 50%. Therefore, in such a situation, nearly all ones would be well-predicted but, in turn, relatively few zeroes will be so²⁹. As this case seems to be close to such situation, the results in the table above should not be surprising. It is interesting, therefore, to gauge the forecasting performance against some other predictor.

Consider a naive predictor of the form

$$\begin{aligned}\hat{d}_h &= 1 \text{ if } \bar{d} > 0.5 \\ \hat{d}_h &= 0 \text{ otherwise;}\end{aligned}$$

(6)

that is, it predicts all ones if the percentage of ones in the sample is greater than 50%. In this case such estimator would correctly predict 66% of the observations in my sample (or approximately a 14% less than the probit). Although the naive predictor (see Greene, 1992) is perhaps too extreme an example, it helps to highlight the fact that this Probit model estimates are not only useful to examine the effect of individual variables on the determinants of the use of vehicles but could also be used to impute use/ownership to data sets containing common explanatory variables but lacking information on the presence of vehicles in the household.

²⁹ Greene (1992) argues that the value of the threshold can be conveniently modified for situations in which some out of sample forecast is needed so that the costs of wrongly classifying observations according to predicted probabilities is minimised.

4 A model for the demand for petrol conditional on vehicle use

The next step consists in studying how households allocate expenditure on petrol. In order to do so I am implicitly assuming a two-stage budgeting framework³⁰. The top stage here consists in allocating total income between leisure, savings and current expenditure. The second stage produces the allocation of such expenditure between petrol and all other goods and this is the process I estimate. The underlying separability assumptions imply that the marginal rate of substitution between goods within the "all other expenditure" group is not affected by the level of consumption of petrol and also that the latter is not affected by changes in the relative prices structure of the former except via income effects.

4.1 Estimation methods

In the last observation for households assumed to be users-owners, recorded zero expenditures correspond exclusively to infrequency of purchase (under the assumptions I have discussed previously). The way to deal with this problem in the context of single equation estimation has been studied by Deaton and Irish (1984), Blundell and Meghir (1987) and Pudney (1989) by means of characterising the likelihood function of data generating processes in which the probability of observing a purchase was explicitly introduced.

However, there exist estimators for infrequency of purchase models which can be implemented in a more straightforward manner outside the realm of maximum likelihood estimation. The first of those is due to Keen (1987) and it starts with the assumption $E(y_{hj}^*) = E(y_{hj})$ alluded to before, that is, expected desired

³⁰ See Deaton and Muellbauer (1980a).

consumption is equal to expected observed expenditures. As equation 2 suggested, this means that the relationship between the two variables can be written in the following way

$$Y_{hj} = D_{hj} \frac{Y_{hj}^*}{P_{hj}} ;$$

$$P(D_{hj}=1) = P_{hj} ;$$
(7)

Where D_{hj} is equal to one if the household is observed purchasing and zero otherwise and P_{hj} is the probability of observing such purchase (recall that this latter probability is conditional on the household being an user/owner). Desired consumption on the j^{th} good can be modelled in the following way

$$Y_{hj}^* = \alpha' z_h + \beta x_h^* + u_{hj} ;$$
(8)

where x_h^* is total desired expenditure and u_{hj} is white noise around desired consumption due to factors other than infrequency of purchase. From equations 7 and 8, we may write

$$Y_{hj} = \frac{D_{hj} (\alpha' z_h + \beta x_h^* + u_{hj})}{P_{hj}} ;$$
(9)

Hence

$$Y_{hj} = \alpha' z_h + \beta x_h^* + \eta_{hj} ;$$

$$\eta_{hj} = \frac{(D_{hj} - P_{hj}) Y_{hj}^* + D_{hj} u_{hj}}{P_{hj}} ;$$
(10)

where the error term η_{hj} is a function of u_{hj} and the factors that determine purchases frequency.

Now note that by the additivity property of demand systems, the following holds for household h 's total observed expenditure

$$\begin{aligned} \sum_j Y_{hj} &= \sum_j (\alpha_j' z_h + \beta_j X_h^* + \eta_{hj}) \\ &= X_h^* + \sum_j \eta_{hj} \end{aligned} \tag{11}$$

The model in equation 10 in terms of observed (as opposite to desired) total expenditures (for estimation purposes) has to be written in the following way

$$Y_{hj} = \alpha' z_h + \beta (\sum_j Y_{hj} - \sum_j \eta_{hj}) + \eta_{hj} \tag{12}$$

and since the term $\sum_j \eta_{hj}$ is not observed, it gets absorbed within a new error term in the following regression equation

$$\begin{aligned} Y_{hj} &= \alpha' z_{hj} + \beta (\sum_j Y_{hj}) + \epsilon_{hj} \\ \epsilon_{hj} &= \eta_{hj} - \beta \sum_j \eta_{hj} \end{aligned} \tag{13}$$

Since this error term is correlated with the total observed expenditure regressor (see eq. 11), ordinary least squares is not a consistent estimator of the parameters of interest. The solution consists in finding a set of instrumental variables, I_h , for the total observed expenditure regressor such that $\text{Cov}(I_h, \sum_j Y_{hj}) \neq 0$ and $\text{Cov}(I_h, \epsilon_{hj}) = 0$. A two-stage least squares estimator that uses all available exogenous variables (including income) to predict total expenditure and thus purge it from the correlation with the error term is therefore a suitable procedure³¹.

This solution is derived for models where total expenditure

³¹ This procedure is in fact a generalised method of moments estimator.

enters the specification of the consumption equation in a linear way (i.e. it has linear Engel Curves, see eq. 8). Therefore, the Instrumental Variable estimator is not consistent when Engel Curves are not linear³². However, it is reasonable to think that there exists a trade off between the degree of flexibility that a model of consumption allows in its Engel curves and the inconsistency bias that one introduces by using Keen's method for its estimation when there exist infrequent purchases. AIM models have been estimated via this method in a considerable amount of literature (e.g. Blundell et al. 1989), so it would seem that the balance is always (rightly) tilted on the side of Engel Curve flexibility.

In this paper I use Keen's estimator to produce one set of estimates for a petrol demand equation on the premise that it pays to do so when the preliminary evidence suggests that Engel Curves are non-linear even at the expense of introducing inconsistency bias. However, I shall also use a recent development in the econometrics literature³³ on infrequency of purchase that provides a consistent estimates when Engel Curves are non-linear. The procedure, which starts from the same basic assumption as Keen's, hinges on the ability to construct, for each household, measures of both desired consumption of petrol and total consumption over all commodities. That is, to purge observed expenditure on petrol and total expenditure from the measurement error induced by infrequent purchases. Another feature of this estimator is that it uses only the households with positive records for all the categories in the demand system³⁴ (under the assumption that there exists independence between the purchase frequency and the consumption equation, this

³² See Hsiao (1989).

³³ Meghir and Robin (1992).

³⁴ In this case these two categories are petrol and the rest of goods. Also, since all households have a positive record for the "rest of goods category" the relevant observations in this application are those with a non-zero record for petrol in the last quarter of observation.

does not produce selectivity bias).

For such households, the equivalent of equation 7 is³⁵

$$Y_{hj}^* = Y_{hj} P_{hj};$$
$$\sum_j Y_{hj}^* = \sum_j Y_{hj} P_{hj};$$

(14)

Thus I can obtain error free measures of both total and petrol consumption provided the category individual probabilities of purchase, the P_{hj} 's, are available. As far as petrol goes, a consistent estimator of this probability for every household is readily available: the number of times a household is observed purchasing petrol out of the total number of times it is interviewed (i.e. the probability of observing a purchase conditional on being an user-owner). Concerning the rest of goods, it is convenient in our circumstances to think in terms of a homogeneous block whose probability of purchase is one (every household does in fact record a positive amount for this category in the sample).

Although I lump together expenditures with different probabilities of purchase in the "rest of goods" group, there is an intuitive justification for the use of observed expenditure on the rest of goods as an exact measure of their underlying consumption: the range of goods included is ample enough to allow category-specific infrequency of purchases to cancel themselves out.

³⁵ There are minor differences in the way Meghir and Robin postulate this relationship: First, they do not use the probabilities of purchase as deflators; instead their "weighting factors" are functions of the number of purchases *within each monitoring period* (such information is not available in the ECPF data). Second, these authors introduce an additional error term reflecting variation around the purchase policy in the link between observed and desired expenditure so that the relationship is given by $Y_{hj}^* = P_{hj}(Y_{hj} - v_{hj})$. In order to develop the estimator, however, they assume that this additional stochastic term is zero (a test for this assumption is also provided by the authors).

Ordinary least squares on the data transformed this way yields consistent estimates of the parameters in the demand equation but the authors recommend to instrument total expenditure with the available exogenous variables to account for its potential endogenous determination.

The functional form I use is the AIM of Deaton and Muellbauer (1980b). The heterogeneity in petrol expenditures and shares alluded to in previous sections could well result in price and income elasticities that vary with demographic characteristics. The AIM lends itself to the inclusion of these effects in a straightforward manner. The model is written in terms of budget shares in the following way

$$w_{hp} = \alpha_p(z_{h1}) + \beta_p(z_{h2}) \log(\sum_j Y_{hj}^* / P) + \gamma_p(z_{h3}) \log p_p + e_{hp} \quad (15)$$

where $\alpha_p(z_{h1})$, $\beta_p(z_{h2})$ and $\gamma_p(z_{h3})$ are linear functions of demographic characteristics, P is a general price index, $\log p_p$ is the logarithm of petrol price relative to a price index of all other goods and e_{hp} picks up the effect of unobservables.

4.2 Estimation results

Tables 6 and 7 present the results for Keens' instrumental variable (IV) and Meghir and Robin re-weighting (RW) estimators respectively.

TABLE 6 PARAMETER ESTIMATES FOR KEEN'S IV ESTIMATOR

VARIABLE	PARAMETER	STANDARD ERROR	T-VALUE
CONSTANT	0.113749	0.029229	3.892
REAL EXPENDITURE AND INTERACTION TERMS			
LOG REAL EXPENDITURE	-0.004665	0.002387	-1.955
" * SELF-EMPLOYMENT	-0.000104	0.000125	-0.834
" WHITE COLLAR EMP.	-0.000098076	0.000210	-0.466
PRICE AND INTERACTION TERMS			
LOG PRICE PETROL	0.043016	0.006211	6.926
" * RURAL DUMMY	-0.004515	0.002980	-1.515
" * CITY DUMMY	0.005435	0.003744	1.452
CONSTANT SHIFTERS			
NUMBER OF MEMBERS	0.002733	0.001205	2.268
NUMBER OF EARNERS	0.001395	0.000819	1.703
CONSTANT SHIFTERS (dummy)			
AGE < 35	0.008829	0.002069	4.267
35 < AGE < 50	0.001724	0.001628	1.059
65 < AGE	0.000080664	0.002348	0.034
UNEMPLOYED	-0.003849	0.002968	-1.297
BLUE COLLAR WORKER	-0.001463	0.001914	-0.764
NOT ACTIVE	-0.008232	0.002147	-3.834
SINGLE PARENT	0.002909	0.013512	0.215
PENSIONER	-0.021614	0.010294	-2.100
ILLITERATE	0.000348	0.001606	0.217
UNIVERSITARY EDUCATION	0.000925	0.002552	0.363
CHILDREN BELOW 6	-0.005361	0.001628	-3.293
CHILDREN BETWEEN 7-14	-0.003856	0.001400	-2.755
CHILDREN BETWEEN 14-23	-0.002686	0.001299	-2.067
OTHER DEPENDENTS	0.002257	0.002743	0.823
HOME RENTER	0.000129	0.001671	0.077
SECOND HOME OWNER	-0.002608	0.001760	-1.482
ADJUSTED R-SQUARED		0.030	
MEAN OF DEPENDENT VARIABLE		0.041	

N.B. Estimating sample are 5107 user-households taken at their last collaboration. Positive records are observed for 3151 households. Method of estimation is 2SLS. The instrument for the real expenditure terms are income, all demographic variables and a series of interactions between the two latter.

TABLE 7 PARAMETER ESTIMATES FOR THE RE-WEIGHING ESTIMATOR

VARIABLE	PARAMETER	STANDARD ERROR	T-VALUE
CONSTANT	0.283271	0.02830	10.01
REAL EXPENDITURE AND DUMMY INTERACTIONS			
LOG REAL EXPENDITURE	-0.016389	0.0023076	-7.10
" * SELF EMPLOYMENT	0.00019552	0.0001150	1.70
" * WHITE COLLAR	0.00012639	0.0001920	0.66
PRICE AND DUMMY INTERACTIONS			
LOG PETROL PRICE	0.032941	0.0059010	5.58
" * RURAL	-0.00952286	0.0028813	-3.31
" * CITY	-0.00001016	0.0036161	-0.00
CONSTANT SHIFTERS			
NUMBER OF MEMBERS	0.00169441	0.0011004	1.54
NUMBER OF EARNERS	0.00072959	0.0007424	0.98
CONSTANT SHIFTERS (dummy)			
AGE < 35	0.00398971	0.0019078	2.09
35 < AGE < 50	0.00222956	0.0015083	1.48
65 < AGE	-0.00054674	0.0023063	-0.24
UNEMPLOYED	-0.00178299	0.0027736	-0.64
BLUE COLLAR WORKER	-0.00026815	0.0017506	-0.15
NOT ACTIVE	-0.00330308	0.0020555	-1.61
SINGLE PARENT	0.016087	0.01786	0.90
PENSIONER	-0.024830	0.01468	-1.69
ILLITERATE	0.00042933	0.0014706	0.29
UNIVERSITY EDUCATION	0.00216129	0.0023469	0.92
CHILDREN BELOW 6	-0.00374149	0.0014889	-2.51
CHILDREN BETWEEN 7-14	-0.00366190	0.0012825	-2.86
CHILDREN BETWEEN 14-23	-0.00301513	0.0011866	-2.54
OTHER DEPENDENTS	0.00370548	0.0025094	1.48
HOME RENTER	0.00292335	0.0015828	1.85
SECOND HOME OWNER	-0.00276018	0.0016274	-1.70
ADJUSTED R-SQUARED		0.122	
MEAN OF DEPENDENT VARIABLE		0.064	

N.B. Estimating sample are 3151 households observed purchasing petrol at their last collaboration period. The dependent variable is deflated by the probability of observing a purchase and total expenditure is similarly adjusted. Method of estimation is 2SLS. The instrument for the real expenditure terms are income, all demographic variables and a series of interactions between the two latter.

4.2.1 Effect of demographic characteristics

The numbers of members and earners seem to have a positive effect on the share although neither effect can be said to be highly significant. On the other hand, if the head of the household is prime aged, the predicted share goes up by a 20% and 5% of the mean for the IV and RW estimators respectively. The effect on the predicted share of being out of the labour force or a pensioner is a reduction of 20% and 50% of the mean respectively with the IV estimator but not significant according to the RW estimate. Children of all ages reduce the predicted share significantly so it seems that either their existence increases total expenditure and therefore petrol shares are lower or there is a shift in the destination of expenditure in detriment of vehicle usage. This is consistent with the idea that if own transport is associated to leisure, the existence of children might reduce the number of outings both because income effects (less money to devote to leisure) and more time taken up by children care.

4.2.2 Expenditure and price effects

The expenditure and price terms in the specification were interacted with occupation and location dummies with the hope of capturing some of the underlying heterogeneity in shares. In particular, I expected the degree of necessity or luxury to vary with the occupation of the head of the household (as self employees might need a vehicle for working purposes and white collar workers may be expected not to use a vehicle as an every day working tool). The results are not supportive of these prior views since the interaction of expenditure terms is not significant in any set of estimates.

One remarkable feature of the results is the difference in magnitude of the estimated parameter on the logarithm of real expenditure between the two models. In particular, the RW estimate is substantially greater in absolute value than its IV counterpart. Both parameters are negative, as one would expect

from the general slope of figure 6.3., but the IV estimator fits a straight line over a cloud of points consisting of positive and zero records. As such zero records are evenly spread over the log of real expenditure axis, the slope of the line (obtained through the IV estimator) remains negative but less pronounced than that of a line fit to the positive observations alone (the RW estimator). The resulting expenditure elasticities evaluated at the mean share are 0.88 and 0.75 for the IV and RW estimators respectively.

As far as prices go, both estimators produce significant coefficients on the log of real petrol price, the RW estimate is smaller in absolute value than the IV one. The IV estimator does not yield very significant interaction responses but the RW estimate of the rural interaction suggests, contrary to prior beliefs, that, at the mean share, petrol demand by rural households is more price elastic, -0.64, than that of the rest of households, -0.5. If it is true that rural households cannot use public transport as much as city dwellers then the higher degree of substitution would suggest that an overall reduction in their (rural households) consumption of transport should take place when the price of petrol increases.

The following table shows mean elasticities by income decile and location as well as the quartile points of the distribution for the RW estimates.

TABLE 8 BREAKDOWN OF ELASTICITIES BY INCOME DECILE AND GEOGRAPHICAL LOCATION AND QUANTILES OF ELASTICITIES DISTRIBUTION FOR THE RE-WEIGHING ESTIMATOR

Expenditure			Price		
rural	inter.	city	rural	inter.	city
.69	.66	.60	-.60	-.41	-.32
.65	.65	.69	-.55	-.39	-.46
.68	.69	.70	-.58	-.46	-.49
.69	.65	.68	-.59	-.39	-.45
.69	.65	.63	-.60	-.40	-.36
.64	.63	.64	-.53	-.37	-.37
.66	.62	.63	-.55	-.35	-.36
.58	.59	.59	-.45	-.29	-.28
.63	.57	.64	-.52	-.25	-.38
.61	.54	.54	-.49	-.21	-.20

Quartile of distribution	5%	25%	50%	75%	95%
Expenditure elasticities	.17	.54	.69	.79	.86
Own price elasticities	-.79	-.66	-.50	-.24	.36

N.B. Elasticities obtained from RW estimates and computed at observed shares for owner households at latest collaboration with positive petrol purchase. The figures in the table are mean elasticities for each decile.

These figures suggest that petrol consumption seems is a necessity for vehicle-owners. The values of the elasticities do not show systematic variation across income deciles or size of township. On the contrary, price elasticities are greater in absolute value for rural households and households at the bottom deciles of income. This effect is common to the two sets of results, although for the IV estimator, the threshold of significance for the price interaction with the rural dummy is above 10%. With respect to the distribution of elasticities, the wide variation is given by the degree of variation in shares themselves. Expenditure elasticities remain always below the unit. However, price elasticities at the top decile are positive because of the existence of very small shares.

5 Possible extensions, conclusions and policy implications

5.1 Methodology

The study of the consumption of petrol with cross section data is specially vulnerable to the influence of unobserved heterogeneity (household fixed effects) through, at least, the following factors: Firstly, the type of vehicle, i.e. motorcycle vs. car or diesel vs. gasoline and engine power. Second, whether the household's type of activity in the labour market requires the use of a vehicle. And last but not least, whether there are public transport services available, how far away is the working place and whether the vehicle is used to travel to and from the latter.

With the ECPF data, it would have been possible, in principle, to implement appropriate estimators to account for these factors³⁶: the within-groups estimator eliminates these effects but still does not address the issue of the measurement error in the total expenditure regressor, which is correlated with the error term via the purchase policy thus violating the strict exogeneity of regressors requirement for the consistency of this estimator. Further, the deviation from time means would be carried out across time periods in which the dependent variable is zero.

Alternative solutions currently under study are: 1) To use households only during the periods in which they record positive expenditures and deflating the latter with the probabilities of purchase (conditional on being an user/owner, of course). This way the problem reduces to a case of measurement error in the total expenditure regressor and, in the context of panel estimators, it can be solved by means of a General Method of

³⁶ In the following discussion I do not contemplate the possibility of individual effects to be uncorrelated with regressors (i.e. the random effects model). Only under exceptional circumstances will the elements in the list of potential sources of individual effects above not be correlated with regressors.

Moments (GMM) estimator that uses lagged values for the error-affected variable in a first differences equation (see Arellano and Bover, 1990). However, this procedure actually needs at least 3 periods and, together with the shrinkage of the sample due to throwing away households during periods of no purchase, the number of observations may become prohibitively small. 2) This suggests that the potentially unique satisfactory solution may be to treat the problem with the maximum likelihood techniques available for limited dependent variables. The method would consist in specifying a likelihood function incorporating the purchase and zero process and their relation to the underlying variables³⁷. In addition it would be necessary to specify a distribution for the individual effect conditional on all leads and lags of the regressors with which it is correlated. The maximisation of such likelihood function on each of the T available cross sections and the corresponding parameter estimates would yield T predictions of the underlying latent variable for each household. To avoid the problem of zeroes in the observed dependent variable, it would then be possible to replace observed expenditures with these predictions so that the parameters of interest can be retrieved by means of a standard within groups estimator (see Arellano and Bover, 1990) or, in order to instrument the (still contaminated) total expenditure regressor, a GMM estimator of the kind mentioned above.

5.2 Main findings

The panel nature of the ECPF can be exploited to eliminate measurement errors that would pose serious difficulties if only cross-section data were available. In the context of private transport demand, the procedure I have used in this paper exploits past information on petrol purchases to indicate the existence of a vehicle in the household thus allowing the distinction between infrequent purchases and the absence of a vehicle. For user-households, I have again taken advantage of the panel in order to compute the individual probabilities of

³⁷ See Blundell and Meghir (1987).

purchase and thus obtain measures of desired consumption of petrol and total expenditure per quarter. This, in turn, permits the consistent estimation of demand equations with non-linear Engel Curves. This type of analysis could, in principle, be useful in studying the demand of problematic categories (such as tobacco, holidays etc) for which the ECPF only provides purchases information.

The paper provides econometric evidence on the determinants of vehicle use. The results suggest that ownership is strongly related to the level of income. As far as distributional implications go, it may be argued that if the passage to ownership by households has been done before 1980 (or there have been second car purchases or replacements of old vehicles for new ones since that year), the increase in progressivity of indirect taxation over the last decade (reported in Mayo and Salas 1993, for instance) can reasonably be linked to a higher rate of vehicles purchases (bearing a 33% purchase tax altogether) amongst households in the top part of the income distribution. The results in this chapter also suggest that rural households may be constrained to consume private transport. They present a higher rate of use-ownership and this could be a consequence of a lower availability of public transport.

The estimated demand models ignore the effect of changes in prices on the stock of vehicles. If expectations are that petrol prices will rise, there would presumably be a reduction in the demand for vehicles and consumers will use more public transport but, by stretching the argument, rich urban dwellers who might perceive an unusually crowded public transport service can then decide to buy a car. As these effects are of second order of magnitude and practically not retrievable from the data, my concern rests on the identification of which households will be hit hardest by petrol taxes. In this sense, the results suggest that the share of petrol out of total expenditure is higher for households in rural zones, where the availability of public transport is presumably lower. A tax on petrol consumption will

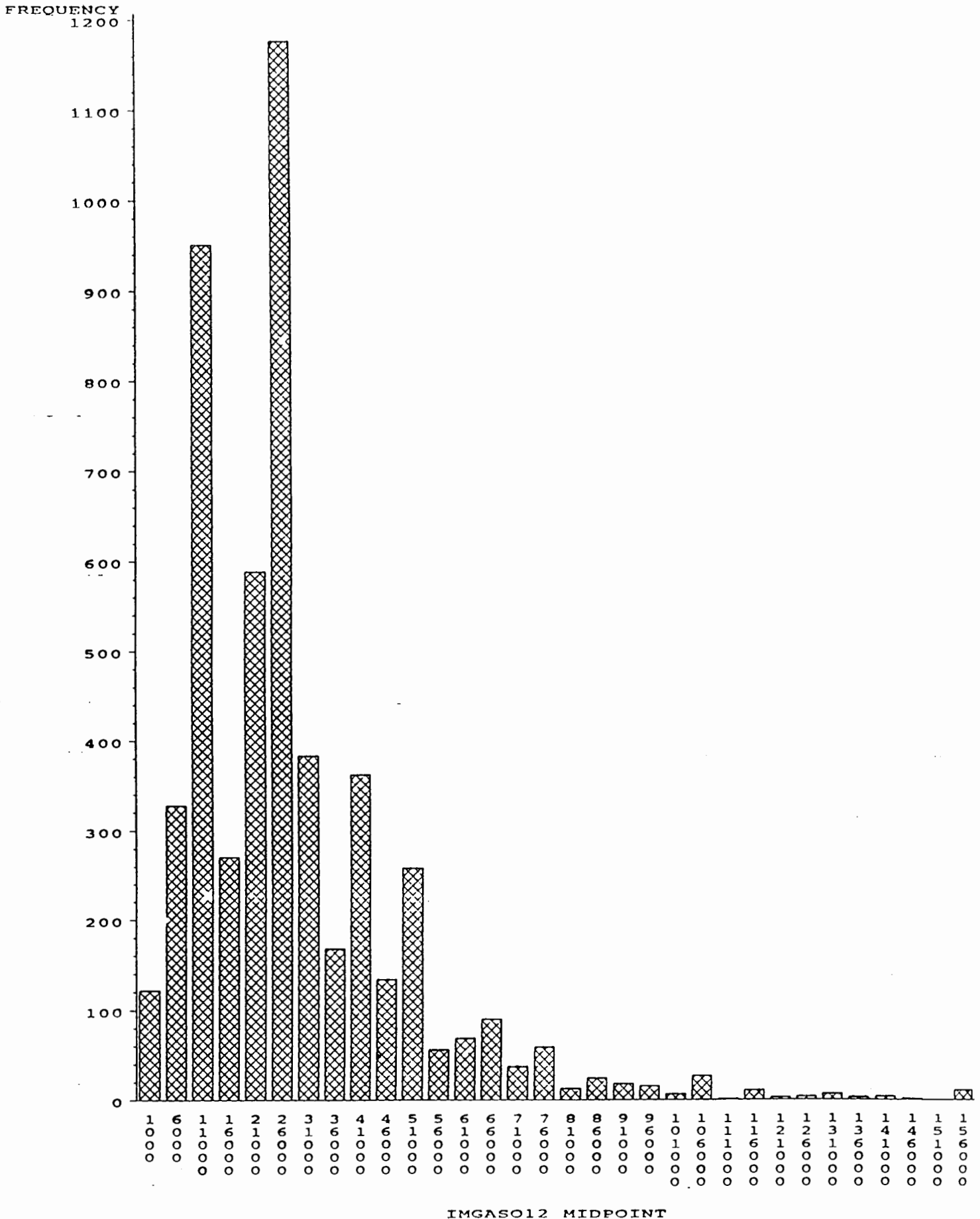
therefore penalise these households much more than those who can easily substitute into public transport.

Price elasticities are higher for rural households according to one set of estimates. This preference for substitution in rural areas means that, whenever public transport is not available, an actual reduction in transport altogether takes place. This does not necessarily affect the welfare of rural households more than that of urban ones. However, on the premise that a high degree of transport is desirable in terms of economic and cultural exchange, this reduction in transport would certainly not bring much benefit to rural settlements.

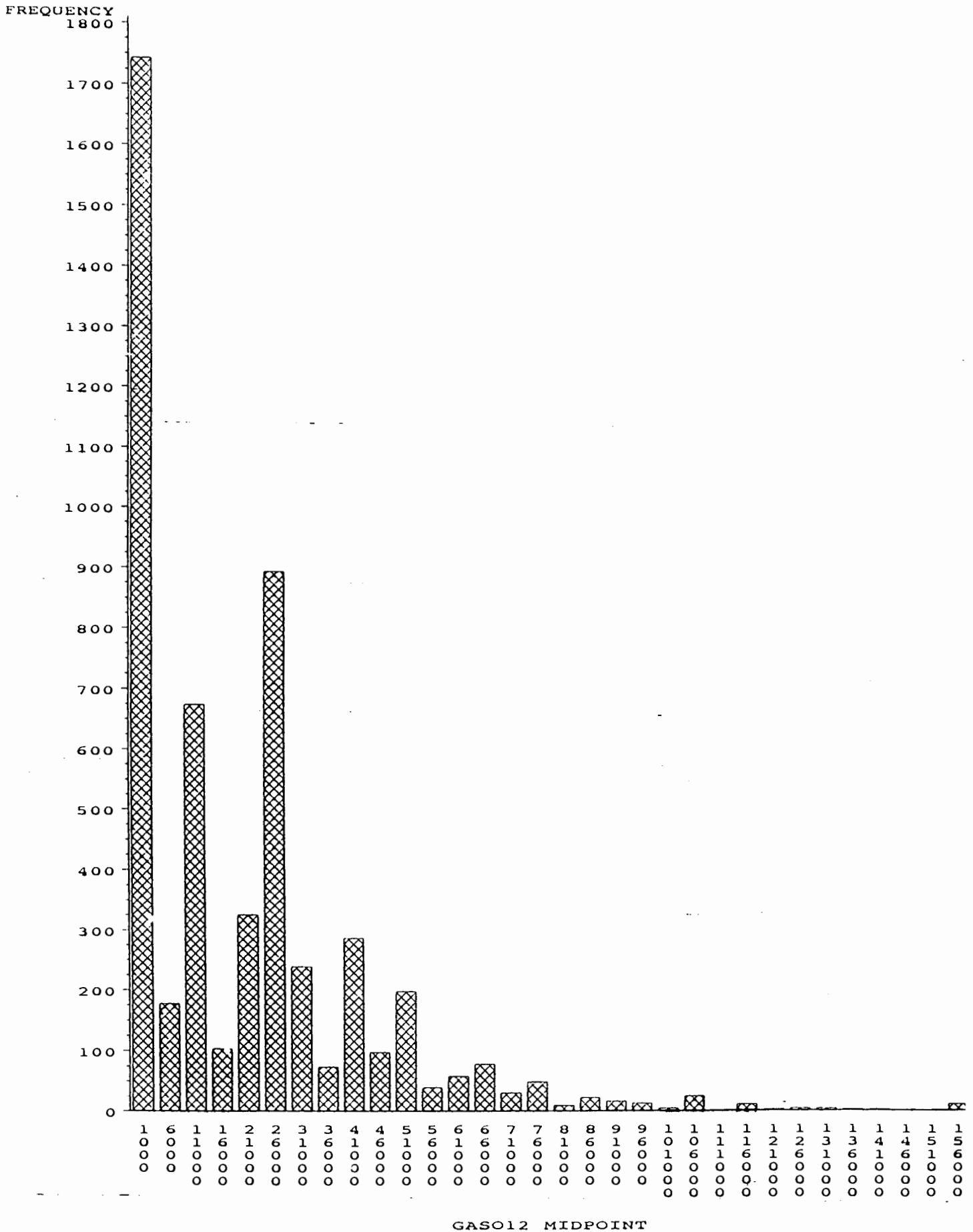
On these grounds urban households and households living in mid-density areas are a more recommendable target for raising revenue through petrol taxes. The scope for levying different rates on private transport is small, as the administrative implementation would not be straightforward and, perhaps most importantly, because the government might only be concerned about raising revenue without much weight being placed on other consequences. But to do so would also be desirable on efficiency and environmental grounds, as big cities in Spain (and elsewhere) suffer well-known problems of dense traffic. To this respect, *The Economist* (6-3-93, p.17) points out: *An ideal solution would be to switch the balance of road-use taxation, away from charges on petrol and car ownership and towards a levy on congestion. (...) Switching the balance to road pricing would make driving cheaper for rural people, but more expensive (though less congested) for city drivers.*

The best way to implement such "desirable" policies in Spain remains an area of interest for future work.

HISTOGRAM 1. DESIRED EXPENDITURE ON PETROL

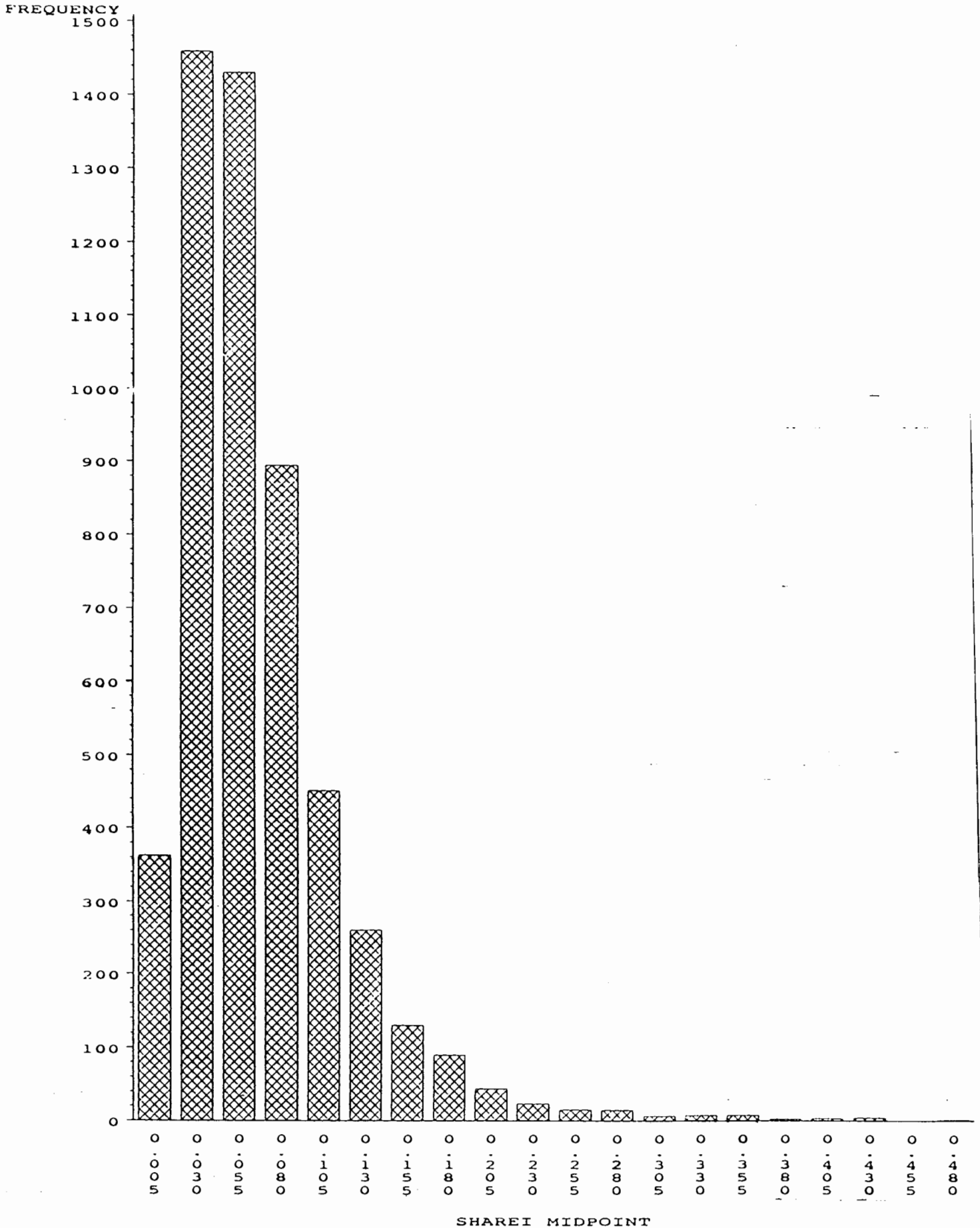


HISTOGRAM 2. OBSERVED EXPENDITURE ON PETROL

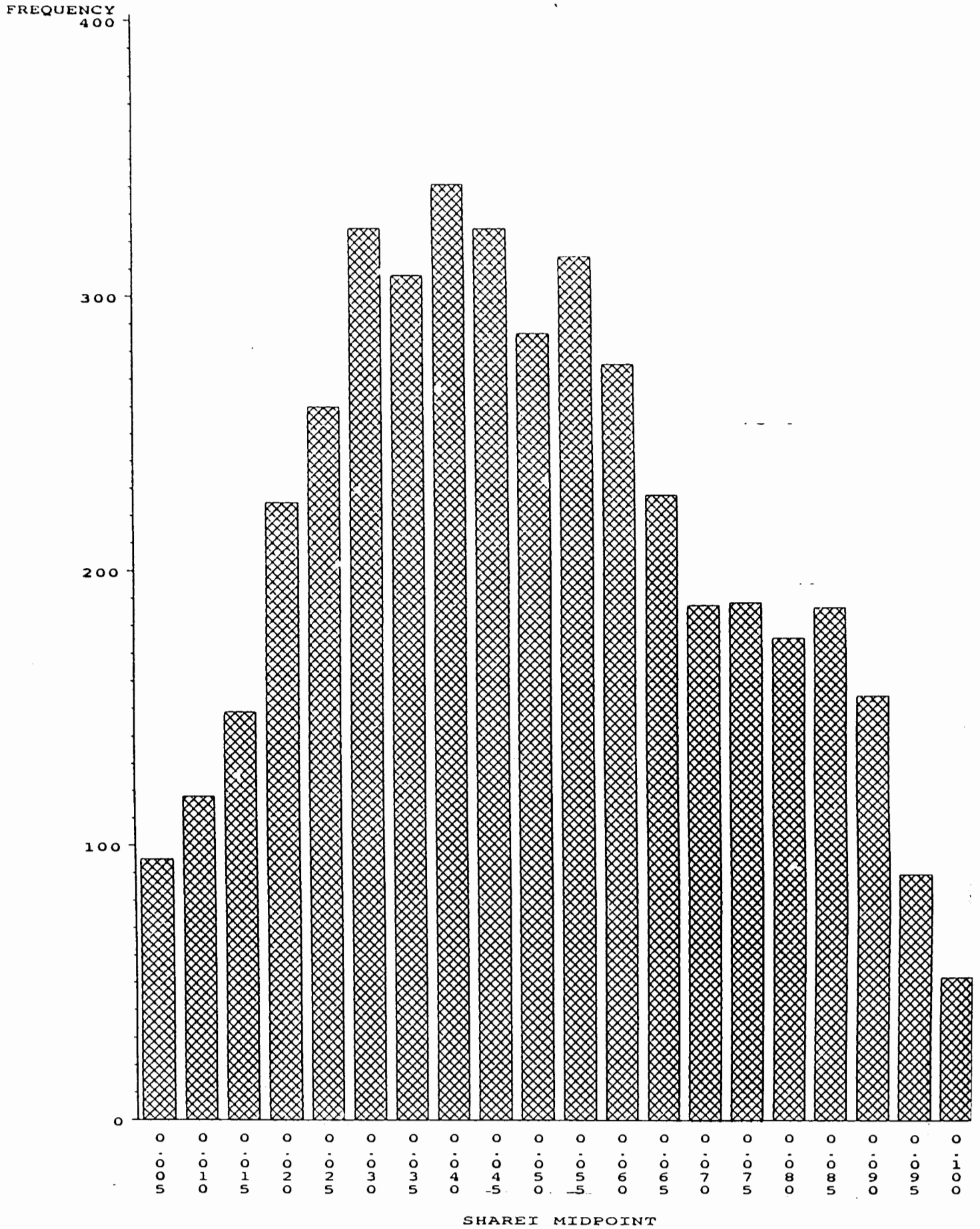


GAS012 MIDPOINT

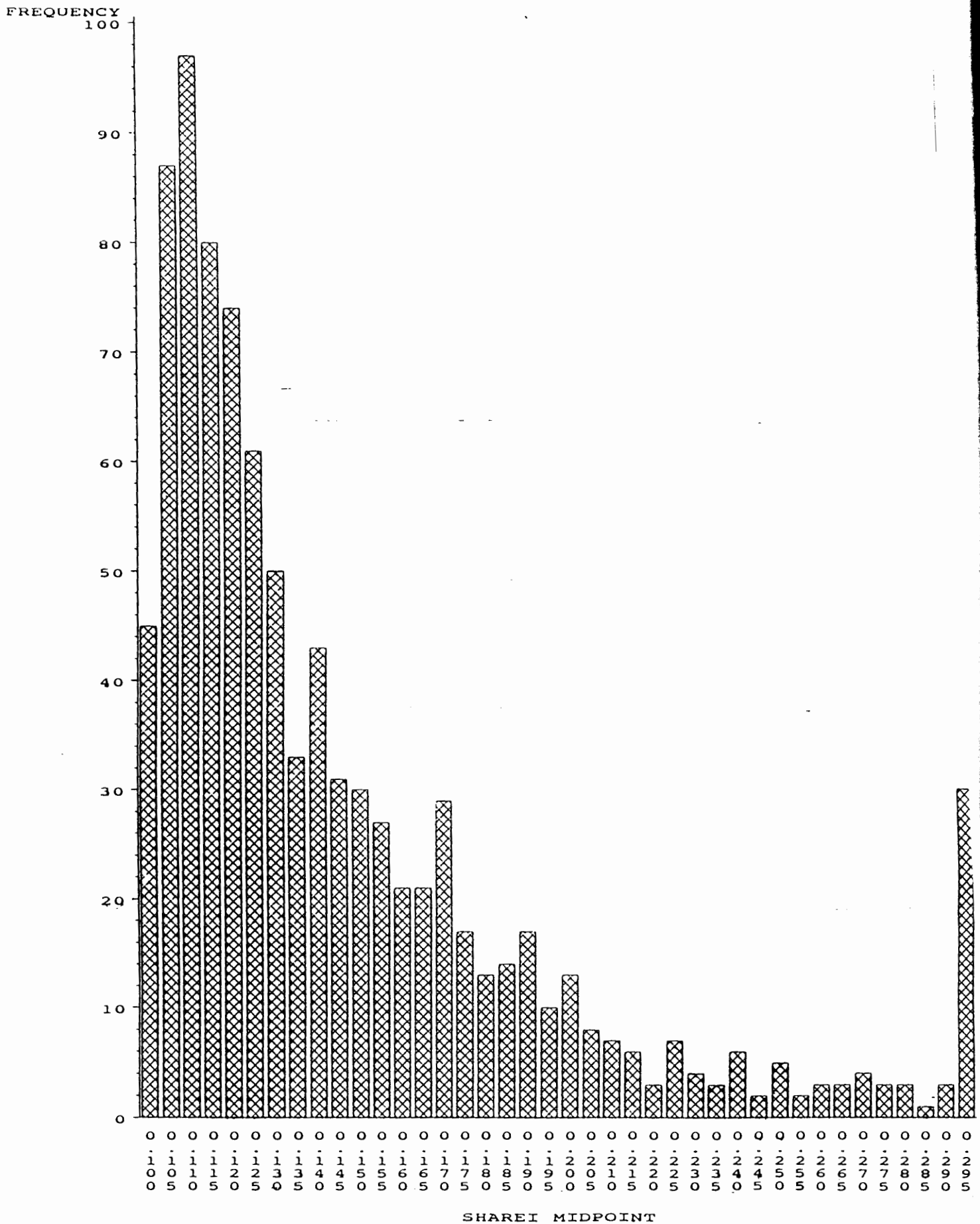
HISTOGRAM 3. ALL "DESIRED" SHARES



HISTOGRAM 4. "DESIRED" SHARES BELOW 10%



HISTOGRAM 5. "DESIRED" SHARES ABOVE 10%



REFERENCES

- Arellano and Meghir (1988). Using complementary data sources. Mimeo.
- Arellano, M. and Bover, O. (1990). "La Econometría de Datos de Panel." Investigaciones Económicas (Segunda Epoca). Vol. XIV no. 1 pp. 3-45.
- Baker, P. Blundell, R and Micklewright, J. (1988). "A cohort Model of Central Heating Ownership in Great Britain." University College London. Department of Economics Discussion Paper no. 88-03.
- Baker, P. Blundell, R and Micklewright, J. (1989). Modelling Household Energy Expenditures Using Microdata. Economic Journal Vol 99 pp 720-38.
- Blundell, R. and Meghir, C. (1987). "Bivariate Alternatives to the Tobit Model." Journal of Econometrics. Vol. 34.
- Blundell, R. Pashardes, P and Weber, G. (1989). "What Do We Learn About Consumer Demand Patterns from Micro-Data?" Institute for Fiscal Studies Micro to Macro Papers No. 3. London.
- Deaton, A. and Irish, M. (1984). "Statistical Models for Zero expenditures in Household Surveys." Journal of Public Economics, vol. 23, pp. 59-80
- Deaton, A. and Muellbauer, J. (1980,a). Economics and Consumer Behaviour. Cambridge University Press.
- Deaton, A. and Muellbauer, J. (1980,b). "An Almost Ideal Demand System." American Economic Review. Vol. 70 No. 3.
- Dubin, J. and McFadden, L. (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. Econometrica Vol. 52, No. 2.
- Greene, W. (1992). Econometric Analysis. McMillan.
- Hsiao, C. (1989). "Consistent estimation of some non-linear in variables errors models." Journal of Econometrics vol. 41, pp. 159-185.
- Johnson, P., McKay, S., Smith, S. (1990). "The distributional consequences of environmental taxes". Commentary no. 23. The Institute for Fiscal Studies. London.
- Keen, M. (1987). "Zero Expenditures and the Estimation of Engel Curves." Journal of Applied Econometrics. No. 1 Vol. 3.

King, M. (1980). An Econometric Model of Tenure Choice and Demand for Housing as a Joint Decision. *Journal of Public Economics* 14 pp. 137-159.

Labeaga, J.M. and López, A. (1992). Simulation of VAT reforms for Spain. Working Paper no. 92/14. Fundación Empresa Pública. Madrid.

Labeaga, J.M. and López, A. (1993). Tax simulations for Spain with a flexible demand system. *Economics working paper* 38. Universitat Pompeu Fabra. Barcelona.

López, A. (1993). Microeconomic models of consumer demand and simulation of indirect tax reforms: an application with Spanish household data. Thesis submitted for the degree of Ph.D. European University Institute. Florence.

Mayo, R. and Salas, R. (1993). "La Progresividad de la Imposición Indirecta. Incidencia de las Pautas de Gasto de los Hogares." *Cuadernos de Actualidad. Hacienda Pública Española* no. 5 pp. 173-184.

Pearson, M. and Smith, S. (1991). "The European Carbon tax: an assessment of the European Commission's proposals". Institute for Fiscal Studies. London.

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

RECENT WORKING PAPERS

1. Albert Marcet and Ramon Marimon
Communication, Commitment and Growth. (June 1991)
[Published in *Journal of Economic Theory* Vol. 58, no. 2, (December 1992)]
2. Antoni Bosch
Economies of Scale, Location, Age and Sex Discrimination in Household Demand. (June 1991)
[Published in *European Economic Review* 35, (1991) 1589-1595]
3. Albert Satorra
Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures. (June 1991)
[Published in *Sociological Methodology* (1992), pp. 249-278, P.V. Marsden Edt. Basil Blackwell: Oxford & Cambridge, MA]
4. Javier Andrés and Jaume Garcia
Wage Determination in the Spanish Industry. (June 1991)
5. Albert Marcet
Solving Non-Linear Stochastic Models by Parameterizing Expectations: An Application to Asset Pricing with Production. (July 1991)
6. Albert Marcet
Simulation Analysis of Dynamic Stochastic Models: Applications to Theory and Estimation. (November 1991), 2d. version (March 1993)
[Forthcoming in *Advances in Econometrics* invited symposia of the Sixth World Congress of the Econometric Society (Eds. JJ. Laffont i C.A. Sims). Cambridge University Press]
7. Xavier Calsamiglia and Alan Kirman
A Unique Informationally Efficient and Decentralized Mechanism with Fair Outcomes. (November 1991)
[Forthcoming in *Econometrica*]
8. Albert Satorra
The Variance Matrix of Sample Second-order Moments in Multivariate Linear Relations. (January 1992)
[Published in *Statistics & Probability Letters* Vol. 15, no. 1, (1992), pp. 63-69]
9. Teresa Garcia-Milà and Therese J. McGuire
Industrial Mix as a Factor in the Growth and Variability of States' Economies. (January 1992)
[Forthcoming in *Regional Science and Urban Economics*]
10. Walter Garcia-Fontes and Hugo Hopenhayn
Entry Restrictions and the Determination of Quality. (February 1992)
11. Guillem López and Adam Robert Wagstaff
Indicadores de Eficiencia en el Sector Hospitalario. (March 1992)
[Published in *Moneda y Crédito* Vol. 196]

12. Daniel Serra and Charles ReVelle
The PQ-Median Problem: Location and Districting of Hierarchical Facilities. Part I (April 1992)
[Published in *Location Science*, Vol. 1, no. 1 (1993)]
13. Daniel Serra and Charles ReVelle
The PQ-Median Problem: Location and Districting of Hierarchical Facilities. Part II: Heuristic Solution Methods. (April 1992)
[Forthcoming in *Location Science*]
14. Juan Pablo Nicolini
Ruling out Speculative Hyperinflations: a Game Theoretic Approach. (April 1992)
15. Albert Marcet and Thomas J. Sargent
Speed of Convergence of Recursive Least Squares Learning with ARMA Perceptions. (May 1992)
[Forthcoming in *Learning and Rationality in Economics*]
16. Albert Satorra
Multi-Sample Analysis of Moment-Structures: Asymptotic Validity of Inferences Based on Second-Order Moments. (June 1992)
[Forthcoming in *Statistical Modelling and Latent Variables* Elsevier, North Holland. K.Haagen, D.J.Bartholomew and M. Deistler (eds.)]
- Special issue** Vernon L. Smith
Experimental Methods in Economics. (June 1992)
17. Albert Marcet and David A. Marshall
Convergence of Approximate Model Solutions to Rational Expectation Equilibria Using the Method of Parameterized Expectations.
18. M. Antònia Monés, Rafael Salas and Eva Ventura
Consumption, Real after Tax Interest Rates and Income Innovations. A Panel Data Analysis. (December 1992)
19. Hugo A. Hopenhayn and Ingrid M. Werner
Information, Liquidity and Asset Trading in a Random Matching Game. (February 1993)
20. Daniel Serra
The Coherent Covering Location Problem. (February 1993)
21. Ramon Marimon, Stephen E. Spear and Shyam Sunder
Expectationally-driven Market Volatility: An Experimental Study. (March 1993)
[Forthcoming in *Journal of Economic Theory*]
22. Giorgia Giovannetti, Albert Marcet and Ramon Marimon
Growth, Capital Flows and Enforcement Constraints: The Case of Africa. (March 1993)
[Published in *European Economic Review* 37, pp. 418-425 (1993)]

23. Ramon Marimon
Adaptive Learning, Evolutionary Dynamics and Equilibrium Selection in Games. (March 1993)
[Published in *European Economic Review* 37 (1993)]
24. Ramon Marimon and Ellen McGrattan
On Adaptive Learning in Strategic Games. (March 1993)
[Forthcoming in *A. Kirman and M. Salmon eds. "Learning and Rationality in Economics"* Basil Blackwell]
25. Ramon Marimon and Shyam Sunder
Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence. (March 1993)
[Forthcoming in *Econometrica*]
26. Jaume Garcia and José M. Labeaga
A Cross-Section Model with Zeros: an Application to the Demand for Tobacco. (March 1993)
27. Xavier Freixas
Short Term Credit Versus Account Receivable Financing. (March 1993)
28. Massimo Motta and George Norman
Does Economic Integration cause Foreign Direct Investment?
(March 1993)
[Published in *Working Paper University of Edinburgh 1993:I*]
29. Jeffrey Prisbrey
An Experimental Analysis of Two-Person Reciprocity Games.
(February 1993)
[Published in *Social Science Working Paper 787* (November 1992)]
30. Hugo A. Hopenhayn and Maria E. Muniagurria
Policy Variability and Economic Growth. (February 1993)
31. Eva Ventura Colera
A Note on Measurement Error and Euler Equations: an Alternative to Log-Linear Approximations. (March 1993)
32. Rafael Crespi Cladera
Protecciones Anti-Opa y Concentración de la Propiedad: el Poder de Voto.
(March 1993)
33. Hugo A. Hopenhayn
The Shakeout. (April 1993)
34. Walter Garcia-Fontes
Price Competition in Segmented Industries. (April 1993)
35. Albert Satorra i Brucart
On the Asymptotic Optimality of Alternative Minimum-Distance Estimators in Linear Latent-Variable Models. (February 1993)

36. Teresa Garcia-Milà, Therese J. McGuire and Robert H. Porter
The Effect of Public Capital in State-Level Production Functions Reconsidered.
(February 1993)
37. Ramon Marimon and Shyam Sunder
Expectations and Learning Under Alternative Monetary Regimes: an Experimental
Approach. (May 1993)
38. José M. Labeaga and Angel López
Tax Simulations for Spain with a Flexible Demand System. (May 1993)
39. Daniel Serra and Charles ReVelle
Market Capture by Two Competitors: The Pre-Emptive Location Problem.
(May 1993)
[Forthcoming in *Journal of Regional Science*]
40. Xavier Cuadras-Morató
Commodity Money in the Presence of Goods of Heterogenous Quality.
(July 1993)
[Forthcoming in *Economic Theory*]
41. M. Antònia Monés and Eva Ventura
Saving Decisions and Fiscal Incentives: A Spanish Panel Based Analysis.
(July 1993)
42. Wouter J. den Haan and Albert Marcet
Accuracy in Simulations. (September 1993)
[Forthcoming in *Review of Economic Studies*]
43. Jordi Galf
Local Externalities, Convex Adjustment Costs and Sunspot Equilibria.
(September 1993)
[Forthcoming in *Journal of Economic Theory*]
44. Jordi Galf
Monopolistic Competition, Endogenous Markups, and Growth. (September 1993)
45. Jordi Galf
Monopolistic Competition, Business Cycles, and the Composition of Aggregate
Demand. (October 1993)
[Forthcoming in *Journal of Economic Theory*]
46. Oriol Amat
The Relationship between Tax Regulations and Financial Accounting: a
Comparison of Germany, Spain and the United Kingdom. (November 1993)
47. Diego Rodríguez and Dimitri Vayanos
Decentralization and the Management of Competition. (November 1993)
48. Diego Rodríguez and Thomas M. Stoker
A Regression Test of Semiparametric Index Model Specification. (November
1993)

49. Oriol Amat and John Blake
Control of the Costs of Quality Management: a Review or Current Practice in Spain. (November 1993)
50. Jeffrey E. Prisbrey
A Bounded Rationality, Evolutionary Model for Behavior in Two Person Reciprocity Games. (November 1993)
51. Lisa Beth Tilis
Economic Applications of Genetic Algorithms as a Markov Process. (November 1993)
52. Ángel López
The Comand for Private Transport in Spain: A Microeconometric Approach. (November 1993)

