DOCUMENT DE TREBALL

CREAP2006-01

ECONOMIC DEVELOPMENT AND CHANGES IN CAR OWNERSHIP PATTERNS

Anna Matas; Josep-LLuís Raymond

CENTRE
DE REFERÈNCIA
EN ECONOMIA APLICADA

ECONOMIC DEVELOPMENT AND CHANGES IN CAR OWNERSHIP PATTERNS*

Anna Matas†
Departament d'Economia Aplicada
Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona (Spain)
anna.matas@uab.es

Josep-LLuís Raymond
Departament d'Economia i Història Econòmica
Universitat Autònoma de Barcelona
08193 Bellaterra, Barcelona (Spain)
josep.raymond@uab.es

Abstract: The contributions of this paper are twofold: On the one hand, the paper analyses the factors determining the growth in car ownership in Spain over the last two decades, and, on the other, the paper provides empirical evidence for a controversial methodological issue. From a methodological point of view, the paper compares the two alternative decision mechanisms used for modelling car ownership: ordered-response versus unordered-response mechanisms. A discrete choice model is estimated at three points in time: 1980, 1990 and 2000. The study concludes that on the basis of forecasting performance, the multinomial logit model and the ordered probit model are almost undistinguishable. As for the empirical results, it can be emphasised that income elasticity is not constant and declines as car ownership increases. Besides, households living in rural areas are less sensitive than those living in urban areas. Car ownership is also sensitive to the quality of public transport for those living in the largest cities. The results also confirmed the existence of a generation effect, which will vanish around the year 2020, a weak life-cycle effect, and a positive effect of employment on the number of cars per household. Finally, the change in the estimated coefficients over time reflects an increase in mobility needs and, consequently, an increase in car ownership.

Keywords: car ownership, ordered probit model, multinomial logit model.

---

* This paper was presented at the European Transport Conference, 2005

† Corresponding author. Tel.: +34.93.581.1578; Fax: +34.93.581.2292

1. Introduction

The central role of cars in industrial countries has led to a large amount of research literature devoted to the analysis of the main determinants of the demand for car ownership. Recent reviews of car ownership literature can be found in Bunch (2000) and De Jong et al. (2004). In recent years, disaggregate models based on individual or household data have become the common approach.

This paper is connected to this literature and has two main objectives. The first is to analyse the factors determining the growth in car ownership in Spain over the last two decades. A discrete choice model on a household level is estimated at three points in time: 1980, 1990 and 2000. On the one hand, this time span allows us to observe how the relationship between car ownership and the explanatory factors has changed from a market in clear expansion to a mature one. On the other, it makes it possible to quantify the contribution of each factor. We use micro data from the Spanish Household Budget Survey (EPF).

The second objective is to provide evidence on empirical grounds about a frequently discussed methodological issue. We compare the two alternative decision mechanisms used for car ownership modelling: ordered-response versus unordered-response mechanisms.

In order to provide some figures for car ownership in Spain, we could say that in 1970 the number of cars per capita was well below the European level, as it corresponded to a less developed country. The expansion of the Spanish economy led to a sharp increase in car demand, so that by the year 2000 Spain had reached 451 cars per 1000 inhabitants, a value close to the European average.

2. The data

2.1. Car ownership data

The study relies on cross-section data from the Spanish Household Surveys for 1980, 1990 and 2000, with sample sizes of 23696, 20927 and 28963 observations respectively. This survey provides a nationally representative sample of the level and structure of family expenditure on a disaggregate level as well as information about the durable goods (including the number of cars) owned by households.

We specify four alternatives for car ownership: zero, one, two, and three or more cars. The car ownership shares and the average level of motorization for the three different years are shown in Table 1. In 1980, only slightly more than 50% of households had access to a car, and only 3.8% owned more than one. Following the economic development in Spain, car ownership levels rose by nearly 70% between 1980 and 2000. In this latter year, the number of families with at least one car was 72.6%, with 17.7% of them owning more than one car. However, these figures do not reflect the total increase in the motorization level of the country. Once we take population growth into account, and hence the number of households, the increase in the number of cars approaches 100%[3].

Changes in car ownership followed a different pattern depending on residential location. The EPF provides information that makes it possible to classify the city of residence according to the level of inhabitants. In this study, we have divided municipalities into three categories: large (those with populations of over half a million), medium (those between 10 000 and 500 000 inhabitants), and small (those with less than 10 000 inhabitants). The size of the municipality can be seen as a proxy for a range of variables affecting car ownership. For instance, different access to public transport or spatial distribution of activities.

[3] From 1980 to 2000, the population of Spain increased by 6.5% and the number of households by 23.4%. These figures reflect a reduction in household size.

Table 2 provides information about car ownership levels for each size of municipality and year. As can be observed, in 1980 the car ownership level was higher in large cities and lower in small ones. However, by 2000 the situation had reversed and families living in large cities had the lowest level of car ownership. Factors explaining such changes are related to the increase in income over the observed period and to different income elasticities for different residential locations and over time. We use an econometric analysis to disentangle each effect.

2.2. Explanatory variables

According to the standard literature, car ownership decisions are related to three classes of variable: the socio-demographic characteristics of the household, the costs of car ownership and use, and residential location.

The socio-economic variables included in the equation are: household income, number of working adults, number of non-working adults, number of children, and the age and sex of the head of the family. The number of children was not statistically significant and was excluded from the final specification. All these variables are provided by the EPF.

Household income can be approximated by two measures, current income and total expenditure. In this study, total expenditure was the preferred option under the assumption that decisions about durables rely more on permanent than on current income, and that total expenditure is a good proxy for permanent income. An additional reason for using total expenditure is that this variable is more reliability recorded than income by the EPF.

Like Bath and Pulugurta (1998), we distinguished between working and non-working adults, the assumption being that a working adult has a greater need for mobility. The age of the head of the household is included in order to take into account the life-cycle effect. Given that this effect is not constant throughout the life cycle, we decided to transform the continuous variable into a discrete one by dividing it into four categories. The decision about the number of categories was based on an initial estimation of the car ownership equation that allowed for a different coefficient for each age. The analysis of these estimated coefficients showed that the life-cycle effect was only clear for the youngest and the oldest heads of household. In accordance with this, the age variable was divided into the following categories: head of household aged under 25 years, aged between 25 and 64 years, aged between 65 and 74 years and aged over 74 years. The effect of gender was gathered by a dummy variable that takes the unit value when the head of the household is male.

The estimation of a car ownership equation with a cross-section sample when all households are faced by the same prices makes it impossible to include a price variable. However, the hedonic price index for cars (which we believe is the relevant one) varies across the three time points (1980, 1990 and 2000). Given the methodological approach used (estimation of an ordered probit model) the effect of price is captured by the constant term in the equations and its variation over time.

In order to take into account the effect of residential location, households were grouped into those living in large, medium or small cities. On the basis of the different observed behaviours of households living in different residential locations, we decided to estimate a separate equation for each size of municipality.

The household car ownership level is also influenced by a generation effect. This effect captures the fact that people born in different generations have not had equal car access. This result is shown in several studies, such as Madré, 1990; Madré and Pirotte, 1997; Dargay and Vythoulkas, 1999 and Dargay, 2001. The estimation of a car ownership equation for three different years makes it possible to test the existence of a generation effect, over and above growth in income and changes in socio-economic variables.

The cohorts were formed by grouping observations in accordance with the date of birth of the head of household. Table 3 illustrates the average car ownership level for 8 different cohorts. A generational trend is shown by the fact that average car ownership is higher for households of the same age but born in more recent generations. This effect is very clear up to the generation born in the forties, and much weaker for more recent generations.

Initially, the 8 cohorts were included in the econometric analysis. Nevertheless, the results showed that the estimated coefficients were not statistically different after the fourth cohort. Therefore, in the final specification only the first 4 cohorts enter the equations, as shown in Table 4.

Finally, for the largest cities it was possible to measure the quality of public transport service. The variable used was vehicle-kilometres run per inhabitant. Table 4 summarises the main descriptive statistics of the variables in the model.

3. The model

3.1. Model formulation

The decision about how many cars are to be owned can be modelled as a discrete choice model, where the alternatives are no cars, one car, two cars, and three or more cars. This decision can be modelled as an ordered-response mechanism or unordered-response mechanism. In an ordered choice model the dependent variable has a natural interpretation as an increasing integer. Specifically, if an ordered response is assumed then the values we assign to each outcome are no longer arbitrary and, moreover, each decision nests the previous one. Car ownership as an ordered process can be found in the works of Kitamura and Bunch, 1990; Pendyala et al., 1995; Dargay and Hanly, 2004 and Giuliano and Dargay, 2006.

However, car ownership models based on unordered response mechanisms have also been estimated in the literature. This is an attractive model specification because it is consistent with the random utility maximization framework. Examples of unordered-response models can be found in Mannering and Winston, 1985; Train, 1986 and Hensher, 1992. Besides, Bath and Pulugurta (1998) compared ordered versus unordered-response mechanisms using several data sets and concluded that the multinomial specification was the preferred option in terms of forecasting and several measures of fit.

In our study, an ordered probit model was the selected specification. The decision was based on two grounds. First, we compared the forecasting performance of an ordered probit model versus a multinomial logit model. The result was, as is detailed below, that from a practical point of view, the models were almost undistinguishable in terms of forecasting capacity. Second, if car ownership is modelled using a multinomial logit, the independence from irrelevant alternatives assumption, which underlines the econometric foundation of the multinomial approach, lacks any meaning. According to the previous reasoning, we consider that, ceteris paribus, the decision must be made on conceptual grounds, and to our judgement, car ownership is better interpreted as an ordered-response mechanism. Moreover, the choice between an ordered probit or ordered logit is an unsolved question. The empirical results of both approaches are very similar. Nonetheless, the normality hypothesis that underlines the probit formulation has a long econometric tradition, which is why we selected this alternative.

The ordered probit model can be derived from a latent variable model. This variable cannot be observed, but does measure the underlying desire for car ownership that can be expressed as

$$y^* = X\beta + \varepsilon \qquad\qquad \varepsilon \sim N(0,1) \qquad\qquad (1)$$

where y^* is the standardized latent variable, X is the set of explanatory variables and ε is the random term.

The observed values for car ownership are related to the latent variable through the following expression:

$$\begin{array}{ll} y = 0 & \text{if } y^* \leq \mu_1 \\ y = 1 & \text{if } \mu_1 < y^* \leq \mu_2 \\ y = 2 & \text{if } \mu_2 < y^* \leq \mu_3 \\ y = 3 & \text{if } y^* > \mu_3 \end{array} \qquad\qquad (2)$$

where μ_1, μ_2, and μ_3 are unknown threshold parameters to be estimated.

Following the methodological approach that goes from general to specific, the estimation strategy consisted of starting from a general model and imposing admissible data constraints until reaching the final constrained specification. So, in order for the general model to nest the constrained models for the different years, the starting model has been formed as a pool that contains the observations for the years 1980, 1990 and 2000. In this general model, the parameters can take different values depending on the year of observation. Therefore, equation (1) can be rewritten as:

$$y_t^* = X_t\beta_t + \varepsilon_t \qquad (4)$$

where t=1980, 1990 and 2000 and y_t^* is the non observable continuous latent variable that expresses the desired degree of motorization for families that are observed in the cross-section corresponding to the year t, X_t is the matrix of explanatory variables and β_t are the parameters to be jointly estimated with μ.

It should also be noted that the estimated model is essentially a static model. As is well known, when the objective of the research is to capture dynamic adjustment the use panel data offers clear advantages over cross-sectional models. This has been stressed for car ownership modelling in several studies (Kitamura and Bunch, 1990; Hensher, 1992; Dargay and Vythoulkas, 1999 and Dargay, 2002). However, given that the purpose of our research is to find out about household behaviour at three different points of time, the static approach will be useful as long as it offers a reliable approximation of the long term relationship. Dargay (2002) offers empirical evidence in favour of this result. Also, preliminary work by the authors supports the same result[4].

[4] An issue frequently discussed in the literature relates to the problems that might arise from the omission of the dynamic structure that underlies the feasible estimates when using a cross section sample. As is well recognized, a behavioral equation that relates a dependent variable with a set of regressors might embody a dynamic structure. However, if a cross section sample is used, the estimable model must omit the dynamic structure because of the lack of statistical information. In the case we were interested in computing, long term responses, the adequacy of the estimation will depend on the extent to which the estimated static model is able to approximate long term responses.

To clarify the previous point, let us assume that the Data Generating Process that relates the dependent variable "Y" to the regressor "X", and taking differences from the average in order to eliminate the constant term, is the following:

$$Y_{it} = \beta_0 \cdot X_{it} + \beta_1 \cdot X_{it-1} + u_{it}$$

where u_{it} is the usual random term, i refers to the individual and t to time period.

Finally, the shift in the threshold parameters "μ" when going from 1980 to 1990, and from 1990 to 2000 captures the influence of all the excluded variables which vary in time but are common for all the households in a given time period. Among others, this is the case for car price.

3.2. Estimation and econometric issues

Estimated results of the general model were agreed with a priori expectations. The differences in the coefficients between the three sample years were clearly significant only for three variables: non-working adults, working adults and the constant term. Hence, we proceeded to simplify the initial model by constraining the coefficients for the rest of the variables to be the same for all three years. The results of the final model specification are shown in Table 5[5].

The dynamic character of the model is reflected in the fact that expected values of the dependent variable depend on the present and past values of the explanatory variable. However, if only information for period "t" is available, the estimated equation will take the following expression:

$$Y_{it} = (\beta_0 + \beta_1)\cdot X_{it} - \beta_1\cdot\Delta X_{it} + u_{it} = k\cdot X_{it} + w_{it}$$

$$w_{it} = (u_{it} - \beta_1\cdot\Delta X_{it})$$

In this case, when estimating the static model using cross section data, the estimated regression coefficient "k" will tend to:

$$\operatorname{limp}(\hat{k}) = k + \frac{\operatorname{limp}(\frac{1}{N}\sum_{i=1}^{N} X_{it}\cdot w_{it})}{\operatorname{limp}(\frac{1}{N}\sum_{i=1}^{N} X_{it}^2)}$$

So, the estimation of the static model will provide a consistent estimation of the long-term response if the ratio of the two probabilistic limits tends to zero. Generally, this will likely occur if the between variation of the explanatory variable (that is, the variation evaluated in the cross section dimension which is captured by: $\frac{1}{N}\sum_{i=1}^{N} X_{it}^2$) is much larger than the within variation (that is, the variation in time for the individual "i" which is captured by: $\frac{1}{N}\sum_{i=1}^{N}\Delta X_{it} = \frac{1}{N}\sum_{i=1}^{N}\left[X_{it} - X_{it-1}\right]$).

In a paper in progress, using the panel structure of the European household budget survey for the period 1994-2001, the authors prove that, in general, using similar formulations to the one employed in this paper (see table 5), the static model estimated with a cross section for a given year tends to offer a good approximation to the long term response of a hypothetical underling dynamic model.

[5] On the basis of Schwarz criteria, the constrained model was preferred to the general model.

A frequently discussed point concerning the estimations that use household budget surveys is whether the observations should be weighted in accordance with their weight in the population. The weighting option implicitly considers that the beta coefficients that we are trying to estimate are not common but specific for each individual. The weighted least squares tries to approximate a pseudo average of the individual beta. However, if the beta coefficients are not common, in general, the weighting option does not necessarily approximate the average of the individual beta. In the case of the usual hypotheses of the regression model being verified, the weighting option leads to a loss of efficiency (Deaton, 1997). In any case, given that in the literature there is no consensus on the most adequate way to proceed, we have estimated the ordered probit models using both alternatives; i.e, weighting the individual observations and not weighting them. Both estimations offered very similar results. Therefore, given the previous considerations, the non-weighting option was preferred.

From this final specification, we proceeded by comparing the forecasting performance of the ordered probit model versus the multinomial logit. In order to evaluate the performance of the two models we created a randomly selected estimation sample containing two thirds of the observations and a validation sample containing one third of the observations not used in estimation. The process was repeated several times.

The forecasting errors for the ordered probit model and multinomial logit were computed in accordance with the following expression:

$e_{ih} = D_{ih} - (P_i)_h$

e_{ih} : Error for household *h* choosing *i* cars

D_{ih} : Dummy variable that takes a unit value if household *h* owns *i* cars and zero otherwise

$(P_i)_h$: Predicted probability of household *h* choosing *i* cars.

The whole process was repeated 5000 times. That made it possible to construct 5000 series of forecasting errors for each level of cars (0, 1, 2 and 3 or more cars). For each series we computed the “mean square error”, defined as the sum of the squares of the forecasting errors divided by the number of forecasted observations. Finally, we analysed the empirical distribution of the mean square errors. To smooth the histogram, a kernel approximation was used.

Table 6 provides a comparison of the distribution function of the mean square error for the multinomial logit and for the ordered probit models in three different ways.

First, we computed the average of the mean square error over the 5000 observations for the multinomial logit and ordered probit models. As it can be observed, the forecasting errors obtained from the two modelling strategies are very similar for all the car alternatives and in the three municipality sizes. Although it is true that in most cases the multinomial logit performs slightly better than the ordered probit, the percentage difference is almost zero. The maximum differences are 2.23% in favour of the multinomial logit for the three car alternatives in large cities and 2.78% in favour of the ordered probit for the three car alternatives in small municipalities.

Second, we computed a 95% confidence interval for the corresponding mean square error. As it can be observed in Table 6, the lower and upper bounds overlap in almost all cases.

Finally, instead of comparing only one of the moments (the mean) or only a part of the distribution (95% confidence interval), and in order to provide a more comprehensive view, a third procedure was designed by comparing the full extent of the distribution. This procedure consists of defining an index that measures the degree of overlapping between the distribution of the mean square error of the multinomial logit and that of the ordered probit. This index will range between one (if the forecasting performance of both models were exactly the same) and zero (if the distributions were completely different)[6]. Figure 1 shows the mean square distributions for zero, one, two and three or more cars corresponding to large municipalities. The blue line shows the kernel that approximates the probability density function of the mean square error for the multinomial logit model and the red line shows the same distribution for the ordered probit model. For each car alternative, we have computed the overlapping area of both

[6] The index is defined as: "Overlapped area/Area under one of the distributions". However, given that in a probability density function the area under the distribution is always one, the index is directly computed as the overlapped area.

distributions. For instance, in the case of large municipalities and zero cars the index is 0.9468. Therefore, the overlapping degree is 94.68%, almost a complete overlap as can be seen in Figure 1. The last column in Table 6 shows this index computed for all car alternatives and municipality sizes.
From the previous empirical analysis, it can be concluded that the forecasting performance of the two modelling strategies are undistinguishable.

4. Estimation results

4.1. Estimated Coefficients

The estimation results for the final specification of the ordered probit model are presented in Table 5. In the three equations, all the estimated coefficients are of the expected sign and generally statistically significant. Moreover, the results agree with available literature (Pendyala et al., 1995; Bath and Pulugurta, 1998 and Dargay, 2002). Although the interpretation of the coefficients of an ordered probit model is not direct, some conclusions can be drawn by comparing the estimated parameters over the three years and different municipality sizes.

The variable with the highest significance level is household total expenditure, as a proxy of permanent income, which enters the equation in logs. It has to be stressed that the estimated coefficient remains stable over time and between different sizes of municipality. In order to illustrate the changes in the relationship between car ownership and income over time and for different residential locations, we computed the average car ownership level with all explanatory variables held at their sample means except for the total expenditure that ranges from 10 000€ to 70 000€. Figure 2 shows the results; in each curve the mean expenditure level for the corresponding sample is marked with an *. It can be observed that in 1980 the relationship between motorisation and income was almost the same for the three sizes of municipality. Differences in average car ownership level between residential locations are explained mainly by differences in household income levels. Ten years later, although similar behaviour is also observed, in the largest cities the level of car ownership was lower for a given level of income.

Finally, in the year 2000, once car ownership was approaching saturation, a different pattern emerges with the car ownership level rising while municipality size decreased. This result suggests greater car dependency in those households living in small towns.

The number of adults in a household has a positive effect on the number of cars owned. This effect is greater for working than for non-working adults. The estimated coefficients are larger in small municipalities compared to large ones, and have been increasing over time. These results reflect the greater mobility needs generated in recent decades, related to the process of residential suburbanization and decentralization of economic, commercial and leisure activities. The probability of owning at least one car also increases when the head of household is a man. Nonetheless, the estimated coefficient is much lower in small cities, where a car can be regarded as a more necessary good.

The estimated coefficients for the age of the head of household shows that the life cycle effect is limited to those below 25 and to those above 65, with a lower probability of owning at least one car in the two extreme age groups. This result suggests that an ageing population might contribute to a reduction in car ownership levels. Nevertheless, the life cycle effect is only clear after 75 years of age, and it is not statistically significant in small municipalities. In this latter case, older households keep their cars, probably due to both higher car dependency and lower maintenance costs (for instance, parking costs). The statistical significance of the cohort variables confirms the existence of a generation effect, which is not explained by the increase in income. However, it is important to point out that this effect vanishes for the generation born in the forties.

In the largest cities, the quality of public transport service proves to be a significant variable. In our view, this is a relevant result as it confirms the effectiveness of a transport policy aimed at reducing car ownership.

Finally, for the large municipalities the two dummy variables that capture the changes in the thresholds between 1980 and 1990 and between 1980 and 2000 are negative and with a low significance level. On the contrary, the coefficients are positive and highly significant for the other two geographical areas. A positive coefficient has to be interpreted as a reduction in the threshold values. Given that car price is one of the main excluded variables, the result will be in accordance with a reduction in the real hedonic prices of cars[7].

4.2. Elasticities and marginal effects

Car ownership elasticities were computed for the two continuous variables in the model: total expenditure (income) and quality of the public transport service. Additionally, we computed the marginal effect on the probability derived from a change in the age structure of heads of household, the extinction of the cohort effect and an increase in the number of working adults.

Income elasticities by municipality size and year are presented in Table 7. Elasticity values correspond to aggregate values for the whole sample and are computed for a unit percentage increase in income. The elasticity for the average car ownership level corresponds to the expected number of cars in a household. Income elasticity is always below unity and decreases over time. The decline in the elasticity value can be explained by income growth that, in turn, implies a higher level of motorization. Therefore, our result, in agreement with Dargay (2001), shows that income elasticity declines as the level of motorization increases and saturation approaches.

Income elasticities present different values and patterns for the three residential location groups. In 1980, the highest elasticity value corresponded to small municipalities, which also had the lowest motorization level. The decline in elasticity over time has been more pronounced for these types of municipality, such that in the year 2000 the highest income elasticity was observed in the large municipalities. For this last year, the estimated elasticities were 0.55, 0.45 and 0.47 for large, medium and small municipalities. These values are similar to those presented in Dargay (2002). This author estimates a long-run elasticity equal to 0.50 in urban areas and 0.35 in rural areas.

[7] According to an estimation of a hedonic price equation by the same authors, the real hedonic price index decreased by 11% between 1980 and 1990 and 21% between 1990 and 2000.

Income elasticities were also computed for the four discrete alternatives: no car, 1 car, 2 cars and 3 or more cars. Essentially, the pattern followed by each one is very similar to that of the average car ownership level, although a steeper decline is observed. It can be noticed that in the year 2000 the probability of owning a car was not sensitive to a change in income. Nevertheless, future income growth will lead to a higher car ownership level given the elasticity value for the second and third car.

Figure 3 shows how the income elasticity of the level of car ownership in the year 2000 declines as income increases for the three sizes of municipality. Elasticity was computed with the variables held at their sample means except for total expenditure that takes values from 12 000 to 45 000€. It can be observed that households living in urban areas with more than 500 000 inhabitants are clearly more elastic than the other two groups. It is also shown that for income levels higher than 30 000€ income elasticity stabilises at around 0.4.

For the large urban areas, it was possible to compute car ownership elasticity with respect to the quality of public transport. The results are presented in Table 8. The elasticity for the average car ownership level is low and decreasing over time. Nonetheless, the quality of public transport has a larger impact on the decision to buy the second or third car. Given that in the recent future car ownership growth will result mainly from an increase in the number of cars per household, these results provide evidence in favour of a transport policy aimed at the improvement of public transport in order to reduce the increase in the number of private cars.

For the discrete variables in the model, we computed the marginal effect on the probability derived from a discrete change in each variable. All the changes refer to the year 2000. Table 9 presents the average car ownership level before the change and the computed marginal effect. First, we computed the marginal impact of the ageing of the Spanish population based on official demographic predictions. That is, the age structure of the heads of household has changed in accordance with those predictions. The results show that the effect of ageing on car ownership is very weak, and almost non-existent for small municipalities. Second, we computed the effect of the removal of those generations with limited access to a private car. In this case, the effect is positive and higher for households living in medium and small municipalities. Finally, the effect of a 10% increase in the number of employed people was also computed. The results show a marginal effect of around 0.02 in the three areas, which implies an aggregate elasticity of around 0.23.

4.3. Contribution of the factors explaining the increase in car ownership

One of the objectives of this paper is to quantify the relative importance of each explanatory factor on the growth in car ownership. To do so we distinguish between the effect derived from a change in the explanatory variables and the effect derived from a change in the estimated coefficients. We have only computed the effect of those variables that show a larger increase over time (total expenditure, quality of public transport and cohorts) and for the coefficients changing over time (constant term, and working and non-working adults). We carried out the computations in accordance with the following simulation procedure:

At period $t+1$ the level of car ownership y_{t+1} is a non-linear function of the estimates β_{t+1} and the values of explanatory variables at $t+1$. That is,

$$y_{t+1} = \phi(\beta_{t+1}, X_{t+1}) = \phi(\beta_t + \Delta\beta, X_t + \Delta X)$$

At year t

$$y_t = \phi(\beta_t, X_t)$$

In order to compute the effect of each variable we proceeded in the following way:

- The effect derived from a change in the explanatory variables was computed as: $\phi(\beta_t, X_t + \Delta X) - \phi(\beta_t, X_t)$
- The effect derived from a change in the estimated coefficients was computed as: $\phi(\beta_t + \Delta\beta, X_t) - \phi(\beta_t, X_t)$
- Finally the joint-mix effect derived from the combined (simultaneous) change in the variables and coefficients is computed as the following difference:

$$\phi(\beta_t + \Delta\beta, X_t + \Delta X) - \phi(\beta_t, X_t) - [\phi(\beta_t, X_t + \Delta X) - \phi(\beta_t, X_t)] -$$
$$-[\phi(\beta_t + \Delta\beta, X_t) - \phi(\beta_t, X_t)]$$

The results are shown in Table 10. The computed joint-mix effects were very small, so they are not detailed in the table.

The results show that explanatory variables have played a different role over time and between municipality sizes. In the first decade, the increase in car ownership in the large urban areas is mainly explained by growth in income (44%) and the change in the coefficients (47%). The positive contribution of the coefficient is the sum of two effects of opposite signs. First, the constant term contributed negatively to car ownership; a possible explanation for this result is that the intercept captures the high increase in the costs of car use in the largest cities, for instance parking and congestion costs. Second, the change in the adult coefficients (both working and non-working) had an enormously positive effect on motorization. In this case, it can be related to the greater mobility needs derived from the suburbanization and decentralization processes that took place in Spain in those years.

In medium and small municipalities, the starting value for motorization was lower, but increased at a higher rate, and factors explaining it are somewhat different. The distinctive features are a lower contribution of income growth and a higher contribution of the generation effect and of the change in the coefficients. In these areas, the intercept term can capture the fall in the real hedonic price of car, which in small municipalities has not been compensated by the increase in the costs of use.

In the second decade, the behavioural pattern of the explanatory variables was more similar between large and medium municipalities. For the first group, it should be emphasised that the improvement in the quality of public transport decreased motorization by 12%. Most of the increase is explained by changes in the coefficients of working and non-working adults, and is hence related to changes in mobility patterns. For medium municipalities, the effect of the constant term was almost zero, in contrast with a positive effect in the previous period, whereas income played a more relevant role. Finally, the factors explaining the growth in car ownership in small municipalities were similar to those for the previous decade.

5. Conclusions

The conclusions of this paper are twofold. On the one hand, they relate to the methodology employed, on the other, to the empirical results.

Methodologically, the study concludes that on the basis of forecasting performance, the multinomial logit model and the ordered probit model are almost undistinguishable. Although the forecasting performance of multinomial logit model in terms of mean square error tends to be slightly better than the forecasting performance of the ordered probit model, the percentage of difference can be considered irrelevant from an applied point of view. Given this result, we argue that the choice between both formulations has to be based more on conceptual arguments than on the capacity of adjustment. In this sense, our choice was the ordered probit model based on the idea that the number of cars has a natural interpretation as an increasing integer.

With regards to the empirical results, first we would like to emphasise that the coefficient estimated for the total expenditure, as a proxy for permanent income, was estimated with a high level of significance and is stable both over time and for the three sizes of municipality. Nonetheless, income elasticity is not constant and declines as the level of car ownership increases. Besides, households living in rural areas are less sensitive to changes in income than those living in urban areas. Car ownership is also sensitive to the quality of public transport for those households living in cities where public transport is a good alternative to private car. Therefore, there is a role for public transport policy aimed at reducing car ownership, specifically the second and third car in the household.

The ageing of the population will have a weak effect on the level of motorization. The statistical significance of the cohort variables confirmed the existence of a generation effect that will result in an automatic increase in car ownership during the next years. This effect, however, will progressively decrease and vanish around the year 2020. In accordance with the results, we can also expect a growth in the level of motorization derived from an increase in the employed population. Related to this, in Spain the increase in the female participation rate will entail greater mobility needs and, consequently, an increase in the number of cars per households.

Finally, the change in the estimated coefficients over time also reflects an increase in mobility needs. The processes of decentralising economic activities and suburbanization add to the number of cars per household and shifts the saturation level.

6. References

Bath, C.R. and Pulugurta, V., 1998. A comparison of two alternatives behavioural choice mechanism for household auto ownership decisions. *Transportation Research B* 32 (1), 61-75.

Bunch, D. S., 2000. Automobile demand and type choice. *Handbook of Transport Modelling* eds. D. A. Hensher and Button, Elsevier, pp. 463-479.

Dargay, J.M., 2001. The effect of income on car ownership: evidence of asymmetry. *Transportation Research Part A*, 35, 807-821.

Dargay, J.M., 2002. Determinants of car ownership in rural an urban areas: a pseudo-panel analysis, *Transportation Research Part E*, 38, 351-366.

Dargay, J.M. and Vythoulkas, P.C., 1999. Estimation of a dynamic car ownership model. *Journal of Transport Economics and Policy*, 33, (3), 287-302.

Dargay, J. M. and Hanly, M., 2004. Volatility of car ownership, commuting and mode and time in the UK. *Proceedings of the World Conference on Transport Research*, Istanbul.

Deaton, A., 1997. *The analysis of household surveys: A microeconometric approach to development policy.* The Johns Hopkings University Press.

Giuliano, G.J. and Dargay, J., 2006. Car Ownership, travel and land use: a comparison of the US and Great Britain, *Transportation Research Part A*, 40, 106-124.

Hensher, D.A.,1992. *Dimensions of automobile demand*, Studies in Regional Science and Urban Economics, 22, Amsterdam, North-Holland.

De Jong, G. et al., 2004. Comparison of car ownership models. *Transport Reviews*, 24, 379-408.

Kitamura, R. and Bunch D.S., 1990. Heterogeneity and state dependence in household car ownership: a panel analysis using ordered-response models with error components. *Transportation and Traffic Theory*, M. Koshi (ed.), Elsevier, pp. 477-496.

Madré, J.L., 1990. Long-term forecasting of car ownership and car use. *Developments in dynamic and activity-based approaches to travel analysis* P. Jones (ed.), Oxford Studies in Transport, Aldershot, Avebury., pp. 406-416.

Madré, J.L and Pirotte, A., 1997. Regionalisation of car-fleet and traffic forecast. *Understanding traffic behaviour in an era of change*, Stopher and Lee-Gosselin (eds.), Pergamon.

Mannering, F. and Winston, C., 1985. A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics*, 16, (2), 215-236.

Pendyala, R.M. et al., 1995. A repeated cross-sectional evaluation of car ownership. *Transportation*, 22, 165-184.

Train, K.E., 1986. *Qualitative choice analysis: Theory, econometrics and application to automobile demand*, MIT Press.

Table 1. Descriptive statistics for car ownership

Cars per household (shares expressed as %)	1980	1990	2000
0	48.8	36.9	27.4
1	47.4	53.1	54.9
2	3.6	9.0	15.5
3 or more	0.2	1.1	2.2
Average car ownership level per household	0.55	0.74	0.92
Increase in car ownership level per household (%)		34.2	24.5
Increase in total number of cars (%)		50.5	42.9

Table 2. Cars per household by size of municipality (shares expressed as percentages)

	Large	Medium	Small
1980			
0	43.75	44.53	59.42
1	51.24	51.18	38.25
2	4.66	4.05	2.19
3	0.35	0.24	0.13
Average	0.62	0.60	0.43
1990			
0	36.80	33.70	43.76
1	53.47	55.65	47.33
2	8.70	9.57	7.90
3	1.03	1.09	1.02
Average	0.74	0.78	0.66
2000			
0	32.05	25.77	28.84
1	52.52	57.16	50.66
2	13.74	15.19	17.23
3	1.69	1.88	3.27
Average	0.85	0.93	0.95

Table 3. Average number of cars per household by cohorts

		1980	1990	2000
Cohort	Date of birth	Nº of cars	nº of cars	nº of cars
1	Before 1910	0.133	0.106	0.229
2	1910-1919	0.304	0.232	0.221
3	1920-1929	0.551	0.544	0.406
4	1930-1939	0.663	0.892	0.801
5	1940-1949	0.787	0.954	1.175
6	1950-1959	0.726	0.945	1.148
7	1960-1969	-	0.857	1.073
8	1970-1979	-	-	0.918

Table 4. Descriptive statistics of the explanatory variables

	Large	Medium	Small
1980			
Total yearly expenditure (€)	27 034	22 939	17 600
Age of the head	48.8	48.3	53.5
Non-working adults	1.25	1.30	1.30
Working adults	1.11	1.08	1.06
Sex (%)	0.83	0.87	0.88
Public transport quality (veh-km/inhabitant)	24.6	-	-
Cohorts (shares as %)			
Before1920	0.226	0.219	0.327
1920-1929	0.219	0.200	0.238
1930-1939	0.228	0.228	0.217
After 1940	0.327	0.353	0.218
Number of observations	3111	13 697	6888
1990			
	Large	Medium	Small
Total yearly expenditure (€)	30 657	25 132	20 294
Afe of the head	52.5	51.2	55.3
Non-working adults	1.29	1.34	1.35
Working adults	1.11	1.07	0.98
Sex (%)	0.79	0.83	0.84
Public transport quality (veh-km/inhabitants)	23.5	-	-
Cohorts (shares as %)			
Before 1920	0.139	0.118	0.177
1920-1929	0.188	0.176	0.231
1930-1939	0.219	0.200	0.206
After 1940	0.454	0.506	0.385
Number of observations	1943	12 994	5990
2000			
	Large	Medium	Small
Total yearly expenditure (€)	33 436	27 619	23 713
Afe of the head	54.9	53.3	56.4
Non-working adults	1.17	1.24	1.22
Working adults	1.12	1.17	1.08
Sex (%)	0.76	0.81	0.84
Public transport quality (veh-km/inhabitant)	33.0	-	-
Cohorts (shares as %)			
Before 1920	0.040	0.036	0.058
1920-1929	0.155	0.128	0.180
1930-1939	0.194	0.185	0.220
After 1940	0.611	0.651	0.542
Number of observations	4206	17 728	7029

Table 5. Estimation results of the Ordered Probit Model

	Large municipalities				**Medium municipalities**				**Small municipalities**			
	Coefficient	Std. Error	z-Statistic	Prob.	Coefficient	Std. Error	z-Statistic	Prob.	Coefficient	Std. Error	z-Statistic	Prob.
Log (total expenditure)	1.14206	0.02893	39.477	0	1.01007	0.01292	78.177	0	1.02247	0.01877	54.482	0
Age < 25	-0.29677	0.11123	-2.668	0.0076	-0.31252	0.05100	-6.128	0	-0.19164	0.09772	-1.961	0.0499
64<Aged<75	-0.10215	0.05251	-1.945	0.0517	-0.05543	0.02463	-2.250	0.0244	-0.03484	0.03427	-1.017	0.3093
Age >=75	-0.23703	0.08652	-2.740	0.0062	-0.24316	0.04001	-6.077	0	-0.05290	0.05353	-0.988	0.323
Non-working adults	0.00829	0.02742	0.302	0.7623	0.00172	0.01337	0.129	0.8977	0.07211	0.02001	3.603	0.0003
Working adults	0.12573	0.03608	3.484	0.0005	0.20679	0.01545	13.389	0	0.24969	0.02102	11.876	0
Gender	0.64382	0.03765	17.099	0	0.53445	0.01820	29.361	0	0.34596	0.03122	11.081	0
Public transport quality	-0.19139	0.02595	-7.374	0	n.a.	n.a.			n.a.	n.a.		
Difference in coefficients 1990												
Constant term	-0.14025	0.09959	-1.408	0.1591	0.09528	0.04137	2.303	0.0213	0.17277	0.06164	2.803	0.0051
Non-working adults	0.07236	0.03950	1.832	0.0669	0.04307	0.01711	2.517	0.0118	0.03206	0.02652	1.209	0.2268
Working adults	0.16013	0.04973	3.220	0.0013	0.10462	0.01992	5.252	0	0.12999	0.02767	4.697	0
Difference in coefficients 2000												
Constant term	-0.15894	0.08436	-1.884	0.0596	0.11345	0.03883	2.921	0.0035	0.36983	0.05863	6.308	0
Non-working adults	0.14619	0.03314	4.411	0	0.10875	0.01607	6.767	0	0.10063	0.02482	4.054	0.0001
Working adults	0.25144	0.04276	5.880	0	0.18250	0.01862	9.803	0	0.20953	0.02642	7.931	0
Cohort 1920-1929	0.21385	0.06477	3.302	0.001	0.21479	0.02956	7.265	0	0.24568	0.03944	6.229	0
Cohort 1930-1939	0.41056	0.07096	5.786	0	0.42082	0.03292	12.782	0	0.41715	0.04511	9.247	0
Cohorts 1940-1980	0.44457	0.07636	5.822	0	0.61309	0.03484	17.599	0	0.66473	0.04924	13.499	0
μ_1	11.55711	0.28524	40.518	0	10.87082	0.12928	84.088	0	11.06762	0.18429	60.057	0
μ_2	13.80800	0.29561	46.710	0	13.16333	0.13379	98.392	0	13.27507	0.19168	69.256	0
μ_3	15.24122	0.30336	50.241	0	14.56323	0.13673	106.510	0	14.63628	0.19588	74.722	0
Observations	9260				44419				19907			
Log likelihood	-6501.556				-31468.11				-13373.92			
Schwarz criterion	1.42395				1.421453				1.353088			
Pseudo-R2	0.285001				0.276245				0.330153			

Table 6. Comparison of distribution of the mean square errors for multinomial logit and ordered probit models

		Confidence interval			Confidence interval		Difference	
	MSE probit	lower limit	upper limit	MSE logit	lower limit	upper limit	in MSE (%)	Common area
Large municipalities								
0 cars	0.14091	0.13487	0.14695	0.14039	0.13424	0.14655	0.37%	94.68%
1 car	0.19432	0.18923	0.19941	0.19292	0.18737	0.19846	0.73%	79.08%
2 cars	0.07251	0.06706	0.07795	0.07243	0.06687	0.07800	0.10%	98.68%
3 cars	0.00975	0.00717	0.01234	0.00954	0.00708	0.01201	2.23%	93.02%
Medium municipalities								
0 cars	0.13256	0.12763	0.13749	0.13195	0.12683	0.13707	0.46%	89.58%
1 car	0.19529	0.19113	0.19945	0.19226	0.18763	0.19689	1.58%	48.36%
2 cars	0.08114	0.07636	0.08593	0.08081	0.07596	0.08566	0.41%	94.68%
3 cars	0.01039	0.00820	0.01257	0.01034	0.00819	0.01250	0.41%	98.60%
Small municipalities								
0 cars	0.13374	0.12982	0.13766	0.13329	0.12919	0.13740	0.33%	91.10%
1 car	0.18497	0.18146	0.18849	0.18328	0.17943	0.18713	0.93%	63.78%
2 cars	0.06731	0.06406	0.07119	0.06763	0.06372	0.07090	-0.47%	92.86%
3 cars	0.01274	0.01124	0.01497	0.01311	0.01088	0.01461	-2.78%	84.72%

MSE: represents the average of the mean square errors for the 5000 observations.

Table 7. Income elasticities of car ownership

		Municipality size		
		Large	Medium	Small
Car ownership				
	1980	0.676	0.606	0.750
	1990	0.590	0.503	0.574
	2000	0.548	0.454	0.468
No car				
	1980	-0.671	-0.601	-0.440
	1990	-0.702	-0.675	-0.548
	2000	-0.757	-0.757	-0.675
1 car				
	1980	0.387	0.377	0.570
	1990	0.203	0.151	0.258
	2000	0.096	0.013	0.003
2 cars				
	1980	1.785	1.666	1.798
	1990	1.412	1.262	1.227
	2000	1.147	0.992	0.808
3 or more cars				
	1980	3.019	2.733	2.773
	1990	2.514	2.259	2.096
	2000	2.176	1.917	1.644

Table 8. Public transport quality elasticities of car ownership

	Car ownership	No car	1 car	2 cars	3 or more cars
1980	-0.112	0.114	-0.069	-0.279	-0.438
1990	-0.098	0.120	-0.038	-0.225	-0.376
2000	-0.091	0.130	-0.020	-0.186	-0.330

Table 9. Impact on car ownership of a change in discrete variables. 2000

	Municipality size					
	Large		Medium		Small	
	Initial car ownership	Change	Initial car ownership	Change	Initial car ownership	Change
Age	0.8498	-0.0097	0.9330	-0.0088	0.9494	-0.0035
Cohorts	0.8498	0.0190	0.9330	0.0416	0.9494	0.0679
Working adults	0.8498	0.0198	0.9330	0.0219	0.9494	0.0220

Table 10. Contribution to the increase of car ownership per household (in percentage)

	Large	Medium	Small
1990/1980			
Average car ownership per household			
1980	0.61532	0.598731	0.42931
1990	0.73961	0.778939	0.66087
Increase	**0.12429**	**0.180208**	**0.23156**
Determinant variables			
Total expenditure	44.3%	19.3%	20.9%
Public transport quality	-4.4%	-	-
Generation effect	13.1%	16.9%	15.5%
Coefficients			
All coefficients	47.4%	58.8%	55.9%
Constant term	-42.0%	19.9%	25.0%
Working and non working	89.5%	38.9%	31.0%
Not explained	-0.5%	5.1%	7.6%
	Large	**Medium**	**Small**
2000/1990			
Average car ownership per household			
1990	0.73961	0.77894	0.660871
2000	0.84978	0.93300	0.949423
Increase	**0.11017**	**0.15407**	**0.288552**
Determinant variables			
Total expenditure	36.1%	25.2%	21.6%
Public transport quality	-11.6%	-	-
Generation effect	15.2%	17.7%	12.4%
Coefficients			
All coefficients	72.3%	54.0%	53.8%
Constant term	-6.7%	4.8%	26.9%
Working and non working	79.0%	49.3%	12.2%
Not explained	-11.9%	3.1%	12.2%

Figure 1. Ordered probit versus multinomial logit
Distribution of the mean square error for large municipalities

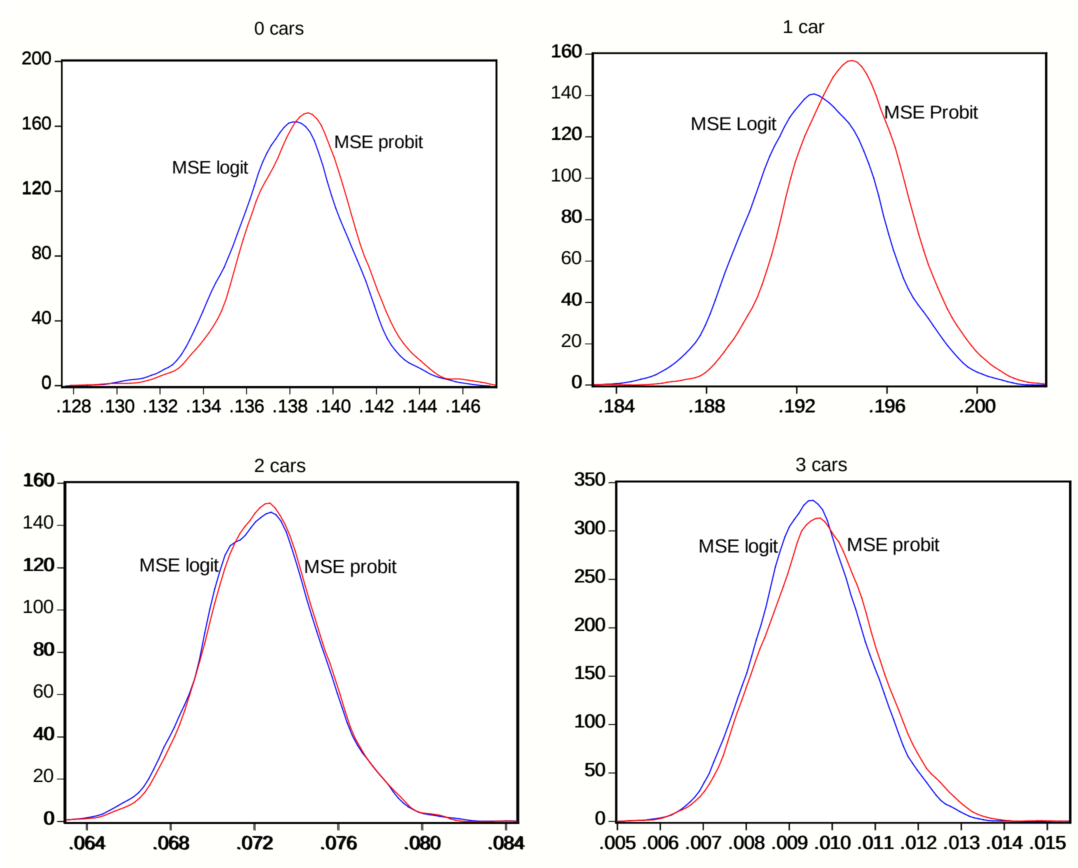


Figure 2. Average car ownership by year and municipality size

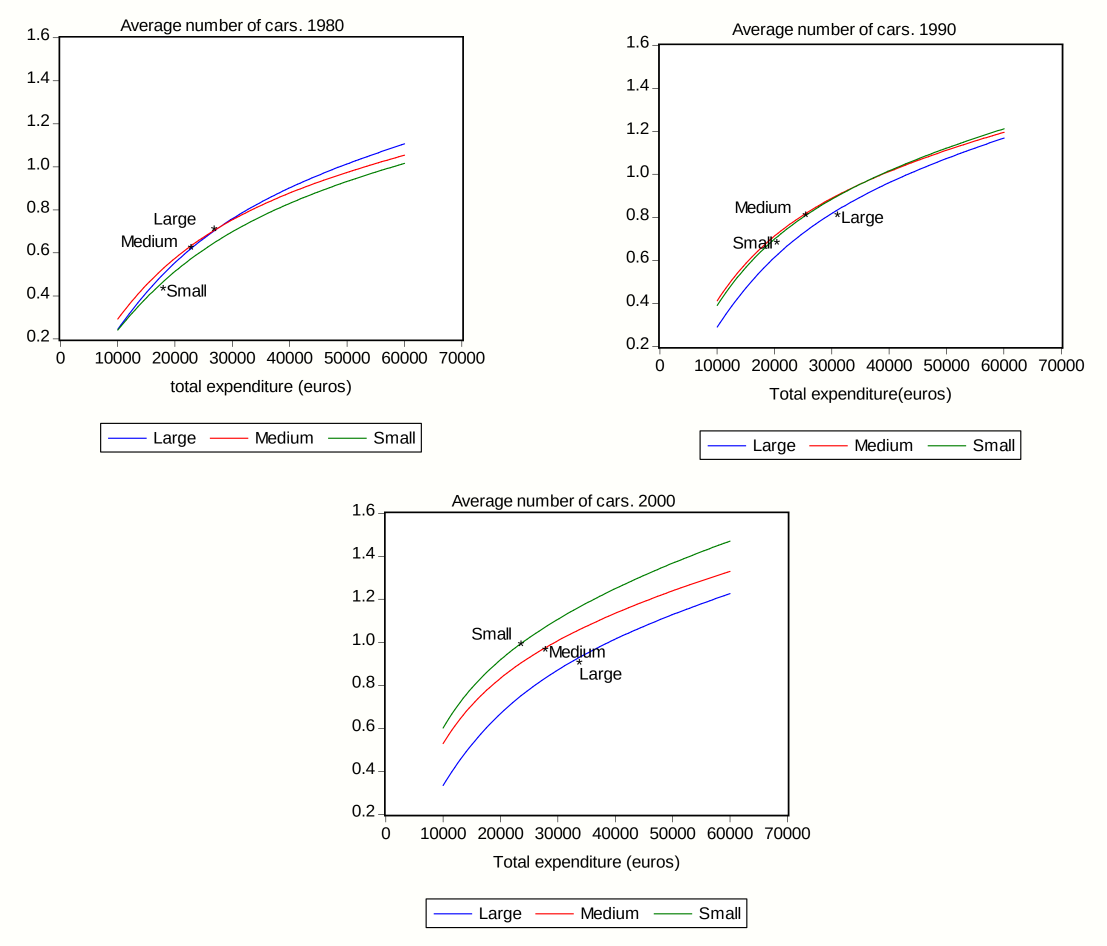
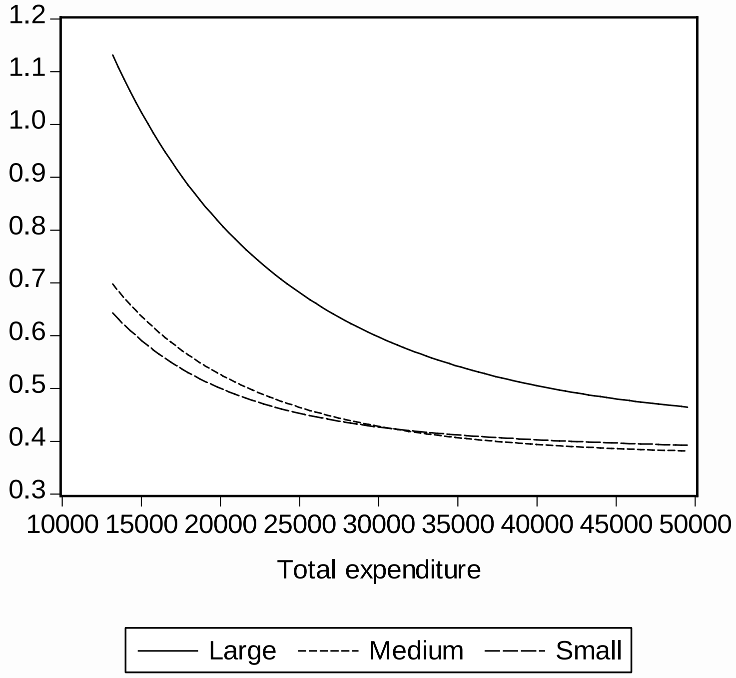


Figure 3. Income elasticity by municipality size



2006

CREAP2006-01
Matas, A. (GEAP)**; Raymond, J.Ll.** (GEAP)
"Economic development and changes in car ownership patterns"
(Juny 2006)

CENTRE DE REFERÈNCIA EN ECONOMIA APLICADA

creap@pcb.ub.es