HOUSING AND URBAN LOCATION DECISIONS IN SPAIN: AN ECONOMETRIC ANALYSIS WITH NON OBSERVED HETEROGENEITY

JOSÉ E. RODRÍGUEZ HERNÁNDEZ

JAVIER A. BARRIOS GARCÍA

Dpto. de Economía Aplicada
Universidad de La Laguna

Dirección de Contacto:

Departamento de Economía Aplicada
Facultad de CC. EE. y Empresariales
UNIVERSIDAD DE LA LAGUNA
Campus de Guajara, s/n. 38297. La Laguna.
Tenerife. Islas Canarias. España
Tel. **(922) 31 70 30; (922) 31 70 28** FAX **(922) 31 72 04**
e-mail: **jerodri@ull.es ; jabarrio@ull.es**

HOUSING AND URBAN LOCATION DECISIONS IN SPAIN: AN ECONOMETRIC ANALYSIS WITH NON OBSERVED HETEROGENEITY

SUMMARY

This paper examines the simultaneous individual choices in Spain as regards housing (tenure choice and demand) and urban location decisions, taking into account certain unobserved heterogeneity that exhibits the behaviour of individuals. With this aim, we propose a multinomial mixed logit model to portray household's decisions considering four different options: home-ownership according to the type of urbanization of the neighbourhood (high level urban property, medium-inferior level urban property, and rural property), and renting. Finally, we estimate housing demand controlling by sample selection. The results obtained permit us to confirm that taking into account both housing tenure and urban location, besides the unobserved heterogeneity behaviour of individuals in this context, could modify appreciably housing demand estimates for both owners and renters.

1. INTRODUCTION

As Goodman (2002) points out, the economics of housing demand has been witness of substantial improvements within the past decades. The availability of better constructed databases, nevertheless generally persisting a large lack with respect to housing items, has helped the joint modelling of tenure choice, mobility and housing demand, using both cross-sectional data

(Börsch-Supan and Pitkin, 1988, Ermisch, 1996, Goodman, 2002) and still short panels (Börsch-Supan and Pollakowski, 1990, Goodman, 2003, Ioannides and Kan, 1996), using enhanced measures of income (current vs permanent), prices (quality adjusted), and other sociodemographic variables.

Recently, some attention has been directed to the differences that may arise in housing demand because the possibility that housing prices were endogenous due to the simultaneous selection of neighbourhood, tenure and housing demand decisions, taking into account the implicit level of public goods, and hence housing prices, provided by each local area. Rapaport (1997) showed that treating housing price as endogenous because of neighbourhood choice results in a sizeable increase in the estimated price elasticity.

This paper is intended to contribute to this discussion in several ways. First, as most of the related literature, we have dealt with a cross sectional database, the Spanish Household Budget Continuous Survey (ECPF) for 1999. We have had to obviate the mobility issue due to lack of data.

Second, we abstract from the "public (and private) goods basket" inherent to a community choice by means of a notion of the degree of urbanization of the neighbourhood. Thus, we initially focus on the simultaneous tenure-urban location decisions understanding these under four different options: homeownership according to the degree of urbanization of the neighbourhood (high level urban property, medium-inferior level urban property, and rural property), and renting. Certainly, it would be desirable to disaggregate the renting decision in the same way, but in Spain the level of renting is under 12% (Housing Census 2001, INE), and hence it is not possible to obtain disaggregated subsamples from our database suitable in size to proceed with a consistent estimation of the model. To explain the tenure-location decision considered, we construct a housing price index tenure and urban location specific, besides considering permanent income and other household characteristics.

Third, we admit the existence of some unobserved heterogeneity in the behaviour of individuals at the time of modelling this discrete choice scenario by using a random parameter multinomial logit (mixed logit) approach (Train, 2003).

Finally, we estimate the demand for both owner-occupied and rented housing conditional on the tenure-urban location decision correcting for sample selection by a generalization of the Heckman's two step procedure (Heckman, 1979) introduced by Barrios (2004). The housing demand equations estimated incorporate the endogeneity of housing prices because they take into account explicitly the tenure-urban location choice, and the housing price is specific for each one of the decision alternatives.

Our findings confirm an inelastic response of housing demand in Spain to income or housing price variations in all the alternatives considered. In particular, we can notice their wide variability among the different tenure-urban location alternatives contemplated, with the larger values for renters, where the special characteristics of the renting sample, possibly more inclined to move than owners due mainly to lower transaction costs, could explain the greater sensitivity to income or housing price shocks.

The organization of the paper is as follows. The next section contains the theoretical and econometric frameworks that we are considering, discussing the previous literature. Section 3 covers sample and variable details, whereas section 4 analyses the results obtained. The last section draws conclusions. At the end, we include two appendixes. Appendix 1 comprises the estimation of the permanent income variable, meanwhile Appendix 2 contemplates the construction of the hedonic housing price indices here used.

2. THEORETICAL AND ECONOMETRIC FRAMEWORK

2.1 Theoretical framework and previous literature

We adopt a theoretical paradigm behind the housing tenure choice problem that is indebted with the seminal contribution of Henderson and Ioannides (1983), and subsequently set up in terms of a dynamic programming problem by Goodman (1995), (2002), Ioannides and Kan (1996) or Gobillon and Le Blanc (2002). These authors model the individual dynamic behaviour confronted simultaneously with the decisions of residential mobility, tenure choice (basically under three possible alternatives: to stay in present housing, to move and own, or to move and rent, exception made of Goodman, 1995, 2002 who does not contemplate explicitly tenure choice) and quantity of housing consumed (and other goods), in presence of transaction costs. The consumer problem consists in finding mobility-tenure decisions and time paths for consuming other goods and housing to maximize the expected utility derived along her living horizon.

From this background follows that the main factors influencing tenure-choice and housing demand, even for those who decide to stay in present housing, are a measure of the income obtained along the individual's living horizon (permanent income), housing and other consumer goods prices, as well as the structure of individual preferences throughout time (that could be determined by certain sociodemographic characteristics).

The empirical contrast more suitable for this theoretical framework should be based on panel data (Börsch-Supan and Pollakowski, 1990, Goodman, 2003, Ioannides and Kan, 1996), but due to the existing lack of panel data containing both adequate information and longitude, a good number of efforts has been directed to test it employing cross sectional data, subsuming the Bellman equations related with the dynamic programming problem considered into an static setting, and using the discrete choice models derived from the random utility maximization tradition (McFadden, 2000) to represent tenure choice or mobility decisions.

In this last situation are the classical papers of Lee and Trost (1978) or Rosen (1979) that analyse simultaneously a binary tenure choice (own vs rent) and housing expenditure using the well known two stage sample selection techniques due originally to Heckman (1979), and based in the probit model, as many others. Catsiapis and Robinson (1982) generalize the Heckman method to address for polychotomous probit, making possible to study jointly tenure-mobility decisions and housing demand by Goodman (2002) or Ermisch (1996). In the same way, Rapaport (1997) analyses housing demand and joint tenure-community choice by an adaptation of Heckman techniques to address for a multinomial logit model at the first step introduced by Dubin and McFadden (1984). In order to represent housing tenure choice alone, Börsch-Supan and Pitkin (1988), Bourassa (1995) or Walker et al. (2002) employ multinomial logit models, and Börsch-Supan and Pitkin (1988) or Skaburskis (1999) utilize the nested logit model distinguishing tenure choice by building type.

Nevertheless, whereas probit, multinomial and nested logit models have been the most widely used to analyse housing decisions, they have some well known limitations, mainly related to computing intricacy in the first case, and to implicit substitution patterns among alternatives in the other two (Train, 2003). In the last years, supported by recent advances in simulation and computing methods, it has been developing the so called mixed logit model (also named random parameter logit or error-components logit) which can address more complex situations, with the important property that any discrete choice model derived from the random utility maximization paradigm has choice probabilities that can be approximated as closely as one pleases by a mixed logit model (McFadden and Train, 2000). The only known previous applications of the mixed logit model to characterize housing tenure choice are Rouwendal and Meijer (2001), Börsch-Supan et al (2001), and Barrios and Rodríguez (2005), being distinguished tenure choice by building type in the last two contributions, with Barrios and Rodríguez (2005) the unique antecedent having dealt with a Spanish sample.

2.2 Econometric modelling

This paper is intended to study jointly housing tenure and urban location decisions. With this aim, we use a mixed logit specification for this discrete choice problem. From this point of view, each individual i (i=1,…,N) is confronted with J alternatives, deriving an utility from the alternative j, U_{ij} determined by:

$$U_{ij}= \beta_{ij}'X_{ij} + \varepsilon_{ij} \qquad i=1,\ldots, N ; j=1,\ldots, J \qquad (1)$$

where $X'=(X_{i1}' \ldots X_{iJ}')\in R^{qxJ}$ is a vector of exogenous observable variables (related to characteristics of individuals and/or alternatives), $\beta_i'=(\beta_{i1}' \ldots \beta_{iJ}')\in R^{q\times J}$ is a random parameter vector that may vary among individuals with joint probability density function $f(\beta|\Omega)$, being Ω the parameters determining this joint density, and $\varepsilon_i'=(\varepsilon_{i1} \ldots \varepsilon_{iJ})$ a random vector with ε_{ij} supposed iid Gumbel (j=1,…,J).

Given β_i, the conditional probability that individual i chooses alternative j, L_{ij}, is determined as in the multinomial logit case:

$$L_{ij}(\beta_i) = \frac{e^{\beta_{ij}'X_{ij}}}{\sum_{k=1}^{J} e^{\beta_{ik}'X_{ik}}} \quad (2)$$

Hence, the unconditional choice probability is given by:

$$P_{ij} = \int_D L_{ij}(\beta)\, f(\beta|\Omega) d\beta \qquad D= \{\beta \in R^{q\times J}\} \quad (3)$$

Because of the multiple integral expression of the choice probabilities, the mixed logit model does not exhibit the IIA property (independence from irrelevant alternatives), and it is usually necessary to adopt simulation techniques for parameter estimation, being commonly used a maximum simulated likelihood procedure with "intelligent" draws (Halton draws) which provide greater accuracy for a given number of draws, shortening computing time (Train, 2003). Under regularity conditions, when the number of draws used rises faster than $\sqrt{N}$, the maximum simulated likelihood estimator is consistent, asymptotically normal and equivalent to the maximum likelihood estimator (Hajivassiliou and Ruud, 1994).

Besides the housing tenure-urban location discrete choice problem, our interest is focused on housing demand conditional on the previous decision. To address the estimation of conditional housing demands, we use a generalization of the Heckman sample selection correction method with a mixed logit at the first step introduced by Barrios (2004).

Following Barrios (2004), suppose that we have an outcome Y_j observed only if the alternative j is chosen, and we wish to model it by an equation of the form:

$$Y_{ij} = \alpha_j' Z_{ij} + v_{ij} \qquad (4)$$

where (omitting the subindex i for simplicity) $Z=(Z_1' \ldots Z_J')$ are observed characteristics of both individuals and alternatives, $\alpha=(\alpha_1' \ldots \alpha_J')$ vectors of unknown parameters, and $v=(v_1 \ldots v_J)'$ represents the impact of unobserved variables. Then, if the following assumptions are fulfilled:

1) v_j is iid on individuals, and $E[v_j| X,Z]=0$, $Var[v_j| X,Z]=\sigma_j^2$, $E[\varepsilon_j| X,Z]=0$, $\forall j=1,...,J$.

2) $E_\beta[\beta_j'X_j]<\infty$, and, $E_\beta[e^{\beta_j'X_j}]<\infty$, $\forall j=1,...,J$.

3) $E[v_j| \varepsilon]$ is linear, $Var[v_j| \varepsilon]$ is constant, and if $\rho_{kj}=corr(\varepsilon_k,v_j)$, then: $\sum_{k=1}^{J}\rho_{kj}^2 <1$, $\forall j=1,...,J$.

Then, the problem of estimation of (4) subject to the mixed logit selection mechanism (1) can be solved consistently in the same way as does the Heckman two stage estimator, that is:

1) We estimate first the mixed logit model (1) to obtain an estimation of $f(\beta|\Omega)$.

2) We can evaluate using simulation methods the artificial variables:

$$\hat{\lambda}_k = E_\beta[\frac{L_k}{1-L_k}\ln L_k],\ k=1,...,J,\ k\neq j,\ \hat{\lambda}_j = E_\beta[\ln L_j], \qquad (5)$$

where L_j, j=1,…,J, denotes the multinomial logit probability of choosing alternative j given in (2). By adding these J artificial variables to the OLS regression (4) we can attain a consistent estimation of the parameters α_j.

It should be noticed that the estimated coefficients of the auxiliary variables $\hat{\lambda}_k$ (k=1,…,J) in the second step, are proportional to the covariance between v_j and ε_k, being positive the constants of proportionality for $k \neq j$, and negative for k=j. As usual, the standard errors of the OLS regression have to be adjusted to account for the first step estimation. To do so, we use an adaptation to maximum simulated likelihood estimation of the method developed by Murphy and Topel (1985, Theorem 2, p. 94).

3. SAMPLE AND VARIABLES

The database used is the Spanish Household Budget Continuous Survey (ECPF) corresponding to 1999. We include in the sample those households that own or rent housing, that is the 98.6% of total, being eliminated other forms of occupancy of dwellings. We also exclude those living in the autonomous cities of Ceuta and Melilla, where the sample size were too small to allow for consistent estimation of hedonic price indices, as well as those who have missing data in any of the variables considered. Thus, although there were 6525 observations in the ECPF sample, only 4460 cases had sufficient information to be usable in this study. The cases finally collected spread relatively evenly within Spain (except Ceuta and Melilla).

Another feature of the ECPF is that the sample is partially updated quarterly, and so, to avoid more reductions in sample size, we have taken the data from the second quarter of 1999, calculating the corresponding annuities for income and housing price variables, these considered in nominal terms.

In order to characterize the choice set on tenure-location considered, first we have tried to define choice alternatives in terms of municipality areas population, obtaining unsatisfactory results pointing out that tenure-location decisions in the sample studied are more concerned on neighbourhoods than on regions or cities, as would be the case if we consider only migrants (Tu and Goldfinch, 1996, Rapaport, 1997). Thus, after taking into account the national coverage of the sample, we have had to abstract from the “public (and private) goods basket” implicit to a neighbourhood choice by means of a notion of degree of urbanization of the local area which involves a ranking of neighbourhoods attending to the type and quality of urbanization prevailing, population or existing commerce, being distinguished finally four housing tenure-urban location alternatives that will be proved statistically relevant in the decision making of the Spanish people:

- *High level urban property*: It includes owning a dwelling in a neighbourhood catalogued as high level urban by ECPF interviewers. This sort of housing includes generally those located in towns over 10.000 inhabitants, in medium-high level neighbourhoods with a smart urbanization, comfortable dwellings, and scarce or medium-high level commerce.

- *Medium-inferior level urban property*: The housing is owned in property and located in a neighbourhood catalogued this way by ECPF interviewers. A medium-lower level urban neighbourhood consists in medium-lower working-class districts, generally with old buildings, popular commerce, as well as depressed areas with scarce urbanization and cheap buildings inhabited in the main by workers without qualification.

- *Rural property*: It contemplates forms of homeownership other than the formers. That is, households living in small towns, generally fewer than 10.000 inhabitants, or in neighbourhoods without urbanization.

- *Renting*: It includes households renting housing. We do not disaggregate this option in the same way as owning because the low level of renting in Spain (under 12% in the last Housing Census 2001, INE) makes not possible to obtain a consistent estimation of the model.

It is noteworthy that in 2000 about 78% of the Spanish population were concentrated in urban areas (Spain Statistical Yearbook 2000, INE), similarly to the urban population level at that year in the US or France, for instance. For a more detailed description of the housing situation in Spain at the provincial level see Barrios and Rodríguez (2004).

We calculate for homeowners the quantity of housing consumed by dividing the self-reported annual market rent (as included in the ECPF) by the regional rent index of a standardized housing in each one of the three owning alternatives considered (regional hedonic housing price index constructed, see Appendix 2). We employ the Duan (1983) “smearing” factor to retransform semi-log estimates of housing price indices in order to avoid retransformation bias. The quantity of standardized units of housing consumed for renters is calculated analogously but using the annual contract rent paid for the dwelling.

In line with the literature concerning housing decisions, we include three types of explanatory variables: sociodemographic characteristics of the household, economic factors, and other characteristics. A brief description of the explanatory variables considered and descriptive statistics of the sample are included in Table 1. The methodological details in the construction of the permanent and transitory income variables, and the hedonic housing price indices used are relegated to appendixes 1 and 2, respectively.

Attending to a classification of regions within Spain according to the housing price level for 1998 based in the mean used and new housing price statistics published by the Ministry of Construction, Housing and National Infrastructures (Ministerio de Fomento), we have defined three dummy variables as explanatory: *Region1*: Regions with mean residential housing prices over the national housing price mean (Baleares, Catalonia, Madrid, Navarra and the Basque country). *Region2*: Regions with mean residential housing prices ranging 80-100% of the national housing price mean (Aragón, Asturias, Canary islands, Cantabria, Castilla-León, Galicia and La Rioja). *Region3* (reference variable): Regions with mean residential housing prices lower than 80% of the national housing price mean (Andalucía, Castilla-la Mancha, C. Valenciana, Extremadura and Murcia).

4. RESULTS

First, we have estimated a mixed logit model for housing tenure-urban location decisions with three random parameters: the coefficients of permanent income in the utility of medium-inferior level urban homeownership and of age in the utility of both high level urban and rural homeownership. We have specified a trivariate normal distribution for these parameters (ie, we admit correlation among the random parameters), whereas the remainder coefficients are considered constants. The selection of this mixing scenario in front of other multiple settings was based on the specification test developed by McFadden and Train (2000).

We used maximum simulated likelihood method to estimate the parameters of the mixed logit model proposed, being employed Nlogit 3.1 software to carry out this procedure. We have considered 200 Halton draws for the simulation process after verifying that above the results were not sensitive to the number of draws adopted. Table 2 reports the conventional multinomial logit as well as the mixed logit estimated models for comparative purposes (we consider renting house as the reference alternative), whereas Table 3 contains the estimates of the Choleski factor associated with the random coefficients. The marginal effects of both models (computed by averaging the individual sample observations) are detailed in Tables 4 and 5.

The results in Table 2 show that the standard deviations of the random parameters in the mixed logit model are all significantly different from zero at a 1% level of significance, pointing out that in fact the influence of both permanent income and head's age variables in this tenure choice model is not perceived in the same way by all the individuals. This outcome is confirmed if we confront the mixed logit and the multinomial models with a log-likelihood ratio test, being rejected the hypothesis that establishes the nullity for the variances of the three random parameters considered. Hence, despite we have included several characteristics of the individuals as explanatory factors in their decision processes, we do note that there is still certain unobserved heterogeneity in their conduct, reflected in the randomness of the permanent income and head's age coefficients.

At the same time, the Choleski factor in Table 3 denotes the existence of correlation among the random parameters of the mixed logit model estimated, being its elements significantly different from zero at a 5% level of significance in the main. This result is corroborated if we compare with a log-likelihood ratio test the same mixed logit model with and without correlation among the random parameters considered.

The main guidelines for the household behaviour regarding housing tenure and location within Spain that can be deducted from the mixed logit model estimated, as summarized in Table 5 averaging by sample observations, are the following:

Permanent income is the most influential variable in the tenure choice-urban location decision. An increase in permanent income implies a rise in the probability of being a home-owner in urban areas, diminishing the probability of owning in rural zones or renting.

At first glance, it could be difficult to analyse the effect of the savings variable because it has as well an indirect influence through the permanent income variable constructed. Nevertheless, as can be deducted from the average direct marginal effect in the mixed logit model, it seems to produce mainly a negative (positive) impact on renting (owning) probability. This way, those individuals who report a greater capacity for saving monthly are more likely to own dwellings.

The age of household head, jointly with housing price and permanent income, have a strong influence on the decision process studied, showing a positive effect on owning alternatives, and a negative one on renting. Therefore, households with a head older are more likely to own.

The head's education does not seem to be very relevant for the housing tenure-urban location decision, although if household head is high school or college-university educated then the probability of owning in a high level urban neighbourhood rises, reducing the likelihood of owning in rural or medium-inferior level urban communities.

Households with a female head tend to become homeowners in urban districts or renting, whereas to be married involves a preference for owning in urban zones, but this factor is not very relevant in the decision making. As well, a greater number of household's members implies a preference for owning dwellings in rural areas and up to a point in high level urban neighbourhoods.

As we expect, in regions within Spain with a higher housing price level, households opt to rent and in a few cases to own within rural areas, at the expense of owning in urban districts.

Finally, it is noteworthy that the constructed (hedonic) housing price variable is highly significant in both estimated models. Moreover, in the mixed logit model this variable turns to be the most relevant factor (with permanent income) to explain housing tenure-urban location decisions, in contrast with the smaller marginal effect derived from the multinomial logit model. The direction of its influence is as expected, that is, rises in housing prices for certain alternative makes to decline the probability of choosing it, increasing all the other probabilities. This fact adds to the empirical evidence accumulated pointing out that the multinomial logit model, as it is proposing a more restricted framework could be underestimating significantly the marginal effects, in our case of permanent income and housing price variables (Hensher, 2001a, 2001b).

Moving on to the housing demand equations, as it is usual in the literature we have specified a log-linear functional form. We include as explanatory variables some of the previously used in the tenure-urban location choice model, specifically, permanent income (in logarithm terms), housing price (hedonic price index in logarithm terms), household head's education level and marital status; and we add a measure of transitory income (in logarithm terms) to reflect that part of the housing demand that can not be explained by permanent income (Goodman and Kawai, 1982, Goodman, 2002, 2003), a dummy variable indicating if the area selected for residence belong to a Spanish provincial capital, where we detect a shifted pattern for housing demand, and four artificial variables denoted λ_1, λ_2, λ_3, λ_4 to correct for sample selection bias. The remainder variables contemplated in the tenure-location choice model estimated previously were introduced at first and did not show a significant effect over the housing quantity demanded, with an F-test rejecting the hypothesis of non null coefficients for these variables.

Table 6 summarizes the housing demand equations estimated for each of the four tenure-urban location regimes distinguished at the first step. The estimation is carried out through conventional OLS corrected for sample selection by the inclusion of the artificial variables λ_1, λ_2, λ_3, λ_4 that were created by simulation using 2000 random draws, after checking the stability of the values obtained above this number of draws. The standard errors for the estimated coefficients in these equations were corrected to take into account the first stage estimation by an adaptation to maximum simulated likelihood of the method developed by Murphy and Topel (1985, Theorem 2, p. 94). For comparative purposes, we include in Table 7 the standard OLS estimates without correcting for sample selection bias.

Looking at Table 6, attention can be drawn to the significance of the estimated coefficients for the artificial variables. Therefore, there is evidence to confirm the existence of sample selection bias and the necessity of taking into account the first step (housing tenure-urban location choice model) to correct for sample selection bias, avoiding inconsistent estimation of model parameters. This fact is confirmed through a log-likelihood ratio test or an F-test contrasting each housing demand model with and without sample selection bias correction (models in Table 6 vs analogous ones in Table 7). At the same time, housing demand equations corrected for sample selection bias present a greater adjusted R^2 than the non corrected ones.

We can have an idea of the quantitative importance of the sample selection bias detected if we compare the corresponding coefficients in Tables 6 and 7. The housing price coefficients in the corrected equations at the three owning alternatives are larger than the corresponding ones in the non corrected equations, especially at the medium-inferior level urban and rural owning. The opposite occurs with the permanent income variable. Even, for renters, the permanent income coefficient at the corrected housing demand is about a 87.4% larger than the non corrected one.

We can highlight the following main consequences from the housing demand equations estimated:

The household head's education variables are significative at a 5% level only for the medium-inferior level urban owning and for renters. The positive sign of their coefficients indicates that a higher level of education rises housing demand in both cases. Similarly, to be married or selecting residence in an Spanish provincial capital also increase housing demand, especially for renters.

Income (permanent and transitory) and housing price variables are again the most influential ones in all the housing demand equations estimated. Following Greene (2000, pp. 928-929) or Goodman (2002, 2003), the conditional housing demand elasticities with respect to housing price and permanent income variables consist of two components: a direct effect on the mean of housing demand through the corresponding coefficient estimated, plus an indirect effect through its presence in the auxiliary variables λ_1, λ_2, λ_3, λ_4 due to its appearance in the probability equations.

Thus, to estimate both permanent income and housing price elasticities, we proceed by increasing 1% each variable for all the sample isolatedly, recalculating the artificial variables, and obtaining the derived percentage change in the housing quantity consumed. By averaging the percentages changes in housing consumption we arrive to an indicator of housing price and permanent income elasticities. Table 8 displays income and price elasticities estimated this way. For comparative purposes, we include as well in Table 8 income and price elasticities estimates for the sample selection corrected model if we adopt at the first step the multinomial logit instead of the mixed logit model (see Table 2) and had applied a sample selection correction method similar to Dubin and McFaden (1984) or Rapaport (1997) (including, as before, the direct effect via the corresponding estimated coefficients and the indirect effect through the auxiliary variables considered).

From Table 8 we can conclude on the one hand that there are little differences between the elasticities estimates taking into account heterogeneity behaviour (ie, considering a mixed logit model at the first step) and obviating this one (ie, with a multinomial logit at the first stage), not appearing a given pattern for the deviations of one model to another. Drawing our attention to the income and housing price elasticities for the housing demand equations corrected with the mixed logit model (left part of Table 8), the models estimated evidence in all the alternatives an inelastic response of housing demand to income or housing price variations. In particular, we can notice their wide variability among the different tenure-urban location alternatives considered, with the larger values for renters, where the special characteristics of the sample, possibly more inclined to move than owners due mainly to lower transaction costs, could explain the greater sensitivity to income or housing price shocks.

Furthermore, we can interpret the coefficients of the auxiliary variables in each one of the housing demand equations obtained in Table 6. For example in the housing demand equation in high level urban neighbourhoods (HU), λ_3 and λ_4 have estimated coefficients 1.927 and 1.218, respectively, both with non null coefficients by a t-test at a 5% level of significance. Thus, we can conclude that non observed factors that increase the probability of being homeowner in a rural area or renting, also increase housing demand in high level urban communities.

For completeness, we have also estimated the standard housing demand equations pooling the sample of owners either, by conventional OLS, and correcting for sample selection bias by the classical Heckman two step method with a probit model at the first stage (including the same explaining variables as the mixed logit model, see Table 2), as applied to housing demand by Lee and Trost (1978) or Rosen (1979). We have considered a housing price for owners corresponding to a medium-inferior level urban neighbourhood, where resides most of the sample studied. Table 9 details housing demand estimates for both owners and renters under these two different scenarios, whereas Table 10 shows the income and housing price elasticities estimates derived, including the direct effect via the coefficient and the indirect effect through the inverse "Mill's ratio" (λ_T) used as auxiliary variable in Heckman's method for sample selection bias correction.

From the results in Table 9, we can confirm the existence of sample selection bias only for renters because the significance of the artificial variable λ_T over a 1% level of significance. If we compare the income and housing price elasticity values from Table 10 to those of Table 8, it is of stress that pooling owners will lead to underestimate notably the housing price elasticity from a value ranging -0.11 to -0.33 for the demand equation corrected by tenure-urban location with a mixed logit model (Table 8) to a value about -0.043 for the pooled sample of owners (Table 10). This confirms the main conclusion in Rapaport (1997). Also, we can observe that permanent income elasticity in the pooled case for owners (0.227) obviate a more sensitive behaviour of rural owners (0.318) as contained in Table 8. Simultaneously, the housing price elasticity for renters in the pooled case is larger than the corresponding one in Table 9, and the permanent income elasticity for renters is even negative in Table 10 due to a large indirect effect through the inverse Mill's ratio (about -1.253 in average).

These results permit us to confirm that taking into account both housing tenure and urban location, besides the unobserved heterogeneity behaviour of individuals in this context, could modify appreciably housing demand estimates for both owners and renters, and thus, obviating these considerations could lead us to distinct housing price and income elasticities estimates, with the corresponding wrong conclusions on the behaviour of households in response to income or housing price shocks.

5. CONCLUSIONS

In this paper we have analysed the decision making process for the Spanish people during 1999 as regards tenure, demand, and urban location of habitual housing, being established four alternatives depending on tenure choice (home owning or renting) and on the urbanization type where it is located the dwelling if it is in property (high level urban, medium-inferior level urban, and rural areas).

Previously, we have constructed a measure of household's permanent income, as well as hedonic price indices for regions in Spain under each one of the alternatives contemplated. This procedure is necessary because this information is not contained in the Spanish Household Budget Continuous Survey (ECPF) which constitutes the data source for this study, and it has revealed as fundamental in the modelization of housing-location decisions.

Afterwards, first we have considered a mixed logit model to stand for this discrete choice scenario. This kind of modelization let us to tackle situations where the choice alternatives present correlation and/or heterocedasticity, allowing to reflect certain non observed heterogeneity in the individuals behaviour by assuming that certain coefficients are random instead of being constants, in our case those of permanent income and household head's age.

The results obtained justify the use of a mixed logit model, being significant the variances of the different random parameters included. This fact confirms that the individuals, confronted with the problem of housing tenure and urban location, does not perceive in the same way variables as permanent income or household head's age.

In addition, it is evidenced that indeed the permanent income and the housing price indices constructed, jointly with the household head's age, turn to be the most relevant factors at the time of explaining households decision making concerning housing tenure and urban location.

Next, we have estimated the demand for both owner-occupied and rented housing conditional on the tenure-urban location decision, correcting for sample selection by a generalization of the Heckman's two step procedure (Heckman, 1979) introduced by Barrios (2004). These estimations incorporates the endogeneity of housing prices because they take into account explicitly the tenure-urban location choice, and the housing price is specific for each one of the decision alternatives.

The housing demand equations estimated show the existence of sample selection bias. Therefore, an estimation of these demand equations without sample selection bias correction would give rise to inconsistent estimators for the demand coefficients.

The more determinant factors in the different housing demand equations are the housing price and household's income. In this sense, the consideration of the mixed logit sample selection correction, and the inclusion of the housing price indices constructed and both permanent and transitory income, allows us to get a better quantification of housing price and income elasticities and to get a better adjustment of the demand equations.

The results obtained evidence an inelastic response of housing demand in Spain to income or housing price variations for all the alternatives. In particular, we can notice their wide variability among the different tenure-urban location alternatives considered, with the larger values for renters, where the special characteristics of the sample, possibly more inclined to move than owners due mainly to lower transaction costs, could explain the greater sensitivity to income or housing price shocks.

The inelastic response observed in all the scenarios with respect to both housing price and income is common to other recent studies on housing demand in Spain. Particularly, Colom *et al* (2002) obtain an income elasticity (disposable income not permanent) of 0.42 and 0.51, and a housing price elasticity of -0.56 and -0.88, for owners and renters respectively. Manrique and Oja (2003) even obviate to include housing price in their housing demand equations obtaining a permanent income elasticity for the demand of primary homes about 0.88 on average. Nevertheless, both studies are based on a sample of the Spanish Household Budget Survey (EPF) conducted between 1990 and 1991, when the macroeconomic context and the housing market behaviour in Spain were substantially different from the situation in the period here analysed. Furthermore, in the last case (Manrique and Oja, 2003) the exclusion of the housing price from the demand equations estimated is probably overvaluing the income effect on housing demand.

It is precisely because we observed different guidelines for the housing demand behaviour of home-owners according to the urban location of the dwelling, we can conclude that to carry out an estimation of housing demand pooling the group of owners will give rise in general to different income and housing price elasticities and, therefore, to different conclusions on the conduct of homeowners after income or housing price shocks.

In fact, the methodology here presented would constitute a powerful tool for urban planners, specially if we restrict our attention to a smaller geographic area, having enough information on neighbourhoods and household characteristics within this area. By developing a mixed logit-OLS framework to display tenure-location choices and housing demand, it is possible to accommodate certain unobserved heterogeneity that use to exhibit households behaviour, and carry out simulation scenarios that can give light to policy makers, besides of course on housing demand forecasting, on urban design, urban regeneration, local housing incentives or subsidies, or other forms of public intervention.

APPENDIX 1: PERMANENT INCOME ESTIMATION

Due to the information available in the ECPF, we follow the yet classical method proposed by Goodman and Kawai (1982) to measure permanent income. Thus, we consider 5259 households data (owners and renters) from the ECPF sample corresponding to 1999 with no missing values in the variables considered, and we proceed to regress the logarithm of the household's annual disposable income in nominal terms (Y_c), as reported by the ECPF, on variables related to household's human and non human capital. The predicted value of this regression provides an estimation of permanent income (in logarithm), meanwhile the residuals are taken as the logarithm of transitory income (Y_T). We include in Table A1 the results of the permanent income regression with t-ratios robusts to heteroscedasticity (White's test).

APPENDIX 2: HEDONIC HOUSING PRICE INDICES

The hedonic housing price indices developed reflects housing price for the different tenure-urban location modes considered. As it is usual, the ECPF does not contain housing price data for each one of the alternatives studied. Therefore, we proceed as, for example, Thibodeau (1995), Ermisch (1996), Goodman (2002), (2003), or Rapaport (1997), to construct hedonic price indices for thirteen regions within Spain in order to approximate the subjective valuation carried by individuals about housing at the different tenure-urban location alternatives.

With this aim, we have estimated, for each one of the thirteen regions, an hedonic house price equation for both, self-reported market rent for owners, and contract rent for renters. The imputed (for owners) and contract (for renters) rents are reported in the ECPF and have been annualized and included as dependent variables in nominal terms.

We include as explanatory variables those reflecting characteristics of the house and of the building where it is located, as well as characteristics of the area, such as population-density or the degree of urbanization. Thus, for the thirteen regions within Spain distinguished, we estimate two separate regressions, each one with the corresponding sample of households in the following way:

$$P_{oi} = \beta_o ' x_{oi} + \mu_{oi} \quad i=1.2....n_o \text{ (owners)} \qquad \text{(A1)}$$

$$P_{rj} = \beta_r ' x_{rj} + \mu_{rj} \quad j=1.2.... n_r \text{ (renters)} \qquad \text{(A2)}$$

Where P_{oi} and P_{rj} are, respectively, the annual rent imputed to owner i (in logarithm) and the annual contract rent (in logarithm) paid by renter j, x_{oi} and x_{rj} are vectors of characteristics of the dwelling and of the area where it is located, β_o and β_r are unknown parameters, and μ_{oi} and μ_{ri} are random terms.

We consider the dependent variables (self-reported and market rents) in logarithm terms because there is some evidence that this transformation has clear advantages on the linear form due to the peculiar characteristics of housing (Malpezzi (2003)).

Following Goodman and Kawai (1982), Ermisch (1996), Rapaport (1997), Rouwendal and Meijer (2001) or Goodman (2002, 2003), once we have estimated the hedonic price equations for different regions within Spain, we have defined a “standard house” departing from the mean values having the different explanatory variables in the full sample. Then, we have calculated over the regions considered, an hedonic price index for the four tenure-location alternatives analysed, helped by the estimated hedonic housing price equations (A1) and (A2) and the standard bundle of housing characteristics previously defined. For the rent alternative, we have used an hedonic price index corresponding to a medium-inferior urban area.

The regressions details should be facilitated upon request to the authors.

Standard housing

Taking into account the mean values for the different housing characteristics included as explanatory in the hedonic housing price equations (A1) and (A2), we have defined a "standard house" as a multifamily dwelling built 30 years ago, with an area of 95 m^2 and 5 rooms, without heating system, and located in a densely populated district not belonging to a Spanish provincial capital. We include in Table A3 the descriptive statistics of the housing characteristics considered.

REFERENCES

BARRIOS, J. A. (2004) Generalized sample selection under RUM, *Economics Letters*, 85, pp.129-132.

BARRIOS, J. A. and RODRÍGUEZ, J. E. (2004) User cost changes, unemployment and home-ownership: evidence from Spain, *Urban Studies*, 41 (3), pp. 563-578.

BARRIOS , J. A. and RODRÍGUEZ, J. E. (2005) Un modelo logit multinomial mixto de tenencia de vivienda en España, *Revista de Economía Aplicada*, forthcoming.

BÖRSCH-SUPAN, A. and PITKIN, J. (1988) On discrete choice models of housing demand, *Journal of Urban Economics* ,24, pp. 153-172.

BÖRSCH-SUPAN, A. and POLLAKOWSKI, H. O. (1990) Estimating housing consumption adjustments from panel data, *Journal of Urban Economics*, 27, pp. 131-150.

BÖRSCH-SUPAN, A., HEISS, F. and SEKO, M. (2001) Housing demand in Germany and Japan, *Journal of Housing Economics*, 10 (3), pp. 229-252.

BOURASSA, S. C. (1995) A model of housing tenure choice in Australia, *Journal of Urban Economics*, 37, pp. 161-175.

CATSIAPIS, G. and ROBINSON, C. (1982) Sample selection bias with multiple selection rules, *Journal of Econometrics*, 18, pp. 351-368.

COLOM, M. C., MARTÍNEZ, R. and MOLÉS, M. C. (2002) Un análisis de las decisiones de formación de hogar, tenencia y demanda de servicios de vivienda de los jóvenes españoles, *Moneda y Crédito*, 215, pp. 199-223.

DUAN, N. (1983) Smearing estimate: a nonparametric retransformation method, *Journal of the American Statistical Association*, 78, pp. 605-610.

DUBIN, J. A. and MCFADDEN, D. L. (1984) An econometric analysis of residential electric appliance holdings and consumption, *Econometrica*, 52, pp. 345-360.

ERMISCH, J. (1996) The demand of housing in Britain and population ageing: microeconometric evidence, *Economica*, 63, pp. 383-404.

GOBILLON, L. and LEBLANC, D. (2002) The impact of borrowing constraints on mobility and tenure choice, *CREST Working Paper 2002-28*.

GOODMAN, A. C. (1995) A dynamic equilibrium model of housing demand and mobility with transaction cost, *Journal of Housing Economics*, 4, pp. 307-327.

GOODMAN, A. C. (2002) Estimating equilibrium housing demand for «stayers», *Journal of Urban Economics*, 51, pp. 1-24.

GOODMAN, A. C. (2003) Following a panel of stayers: length of stay, tenure choice and housing demand, *Journal of Housing Economics*, 12, pp. 106-133.

GOODMAN, A. C. and KAWAI, M. (1982) Permanent income, hedonic price, and demand for housing: new evidence, *Journal of Urban Economics*, 12, pp. 214-237.

GOODMAN, A. C. and THIBODEAU, T. G. (1998), Housing market segmentation, *Journal of housing economics*, 7, pp. 121-143.

GREENE, W. H. (2000) *Econometric Analysis,* 4th edition, Prentice-Hall, Upper Saddle River, NJ.

HAJIVASSILIOU, V. and RUUD, P. (1994) Classical estimation methods for LDV models using simulation, in R. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics Vol. IV*, Elsevier, Nueva York, pp. 2383-2441.

HECKMAN, J. J. (1979) Sample selection bias as a specification error, *Econometrica*, 47 (1), pp. 153-161.

HENDERSON, J. V. and IOANNIDES, Y. M. (1983) A model of housing tenure choice, *American Economic Review*, 73 (1), pp. 98-113.

HENSHER, D. A. (2001) The sensitivity of the valuation of travel time savings to the specification of unobserved effects, *Transportation Research Part E* 37, pp. 129-142.

HENSHER, D. A. (2001) The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications, *Transportation* 28, pp. 101-118.

IOANNIDES, Y. M. and KAN, K. (1996) Structural estimation of residential mobility and housing tenure choice, *Journal of Regional Sci*ence, 36 (3), pp. 335-363.

LEE, L. F. and TROST, R. P. (1978) Estimation of some limits dependent variable models with application to housing demand, *Journal of Econometrics*, 8, pp. 357-382.

MALPEZZI, S. (2003) Hedonic pricing models: a selective and applied review, in A. O´Sullivan and K. Gibb (Eds), *Housing economics and public policy*, Blackwell Publishers, Oxford, pp. 67-89.

MCFADDEN, D. L. (2000) Disaggregate behavioural travel demand's RUM side: a 30-year retrospective, Conference at the International Association of Travel Behavior Analyst, Brisbane, Australia, July 2-7.

MCFADDEN, D. L. and TRAIN, K. (2000) Mixed MNL models for discrete response, *Journal of Applied Econometrics*, 15 (5), pp. 447-470.

MANRIQUE, J. and OJAH, K. (2003) The demand for housing in Spain: an endogenous switching regression analysis, *Applied Economics*, 35, pp. 323-336.

MURPHY, K. M. and TOPEL, R. H. (1985) Estimation and inference in two-step econometric models, *Journal of Business and Economic Statistics*, 3 (4), pp. 88-97.

RAPAPORT, C. (1997) Housing demand and community choice: an empirical analysis, *Journal of Urban Economics*, 42, pp. 243-260.

ROSEN, H. S. (1979) Housing decisions and the U.S. income tax: an econometric analysis, *Journal of Public Economics*, 11, pp. 1-23.

ROUWENDAL, J. and MEIJER, E. (2001) Preferences for housing, jobs, and commuting: a mixed logit analysis, *Journal of Regional Science*, 41(3), pp. 475-505.

SKABURSKIS, A. (1999) Modelling the choice of tenure and building type, *Urban Studies*, 36 (13), pp. 2199-2215.

THIBODEAU, T. G. (1995) House price indices from the 1984-1992 MSA American housing surveys, *Journal of Housing Research*, 6, pp. 439-481.

TRAIN, K. (2003) *Discrete choice methods with simulation*, Cambridge University Press, Cambridge.

TU, Y. and GOLDFINCH, J. (1996) A two-stage housing choice forecasting model, *Urban Studies*, 33 (3), pp. 517-537.

WALKER, B. MARSH, A., WARDMAN, M. and NINER, P. (2002) Modelling tenants' choices in the public rented sector: a stated preference approach, *Urban Studies*, 39 (4), pp. 665-688.

TABLE 1. List of explanatory variables and descriptive statistics

Variables	Means	Standard deviations	Brief description
Sociodemographic factors			
Household size	3.188	1.303	Number of household's members including household head.
Head's age	53.913	14.659	Age of household head and its square value.
Head's age2	3121.390	1636.876	
Male head	0.849	0.358	Male household head=1; otherwise=0.
Head's education1♦	0.524	0.499	Three dummies related to the highest education level attained by the household head: *Head's education1*= primary schooling or lower, *Head's education2*= high school education, *Head's education3*= College-university education
Head's education2	0.299	0.458	
Head's education3	0.175	0.380	
Married	0.802	0.399	Marital status of household's head: married=1, otherwise=0.
Savings	0.376	0.484	A dummy variable reflecting the household possibility of saving monthly.
Economic factors			
*Current income**	14.550	0.584	Current disposable income
*Permanent income**	14.557	0.437	Imputed permanent income
*Transitory income**	-0.006	0.380	Imputed transitory income
*HU owning price**	13.417	0.248	Housing price indices for the different tenure-urban location modes considered.
*MIU owning price**	13.200	0.173	
*Rural owning price**	13.173	0.241	
*Renting price**	12.740	0.480	
Other characteristics			
Region1	0.350	0.477	Three dummies reflecting a classification of regions within Spain attending to the mean housing price level for 1998
Region2	0.317	0.465	
Region 3♦	0.333	0.471	
Provincial capital	0.424	0.494	A dummy variable indicating if the dwelling is located in a Spanish provincial capital

	Sample size	% total sample
HU owners	323	7.2%
MIU owners	2830	63.5%
Rural owners	803	18%
Renters	504	11.3%
Total	4460	100%

Note: ♦ Reference variable.
* Variable in logarithm terms.

TABLE 2. Estimates of multinomial and mixed logit models of housing tenure-urban location

Variables	Multinomial Logit Parameters	Multinomial Logit t-statistics	Mixed Logit Parameters	Mixed Logit t-statistics
Homeownership in high level urban area (HHU)				
Constant	-25.3620	-6.255	-36.5220	-5.163
Household size	0.0466	0.604	0.0595	0.559
Head's age: Mean	0.0844	2.221	0.2434	2.962
Std. deviation			0.0565	3.451
Head's age2	-0.0002	-0.706	-0.0013	-1.933
Male head	-0.0908	-0.330	0.4112	0.970
Head's education2	1.3527	5.668	1.5012	4.645
Head's education3	1.8487	6.640	1.5130	3.365
Region1	-1.3120	-6.017	-2.2423	-5.366
Region2	-1.3374	-6.463	-2.0152	-5.850
Married	0.8179	2.811	1.2650	3.053
Savings	0.9603	5.113	1.4656	4.452
Perm. income	1.3823	4.644	1.9573	4.093
Homeownership in medium-inferior level urban area (HMIU)				
Constant	-8.6767	-3.302	-19.3690	-3.214
Household size	0.0029	0.057	0.0239	0.266
Head's age	0.1035	4.435	0.2300	3.714
Head's age2	-0.0007	-3.548	-0.0015	-3.081
Male head	0.2465	1.491	0.7489	2.006
Head's education2	-0.0876	-0.633	-0.0096	-0.039
Head's education3	-0.6370	-3.456	-1.1366	-2.641
Region1	-1.0087	-6.572	-1.8789	-5.009
Region2	-0.8716	-6.291	-1.5578	-4.728
Married	0.6423	3.797	1.1331	3.320
Savings	0.7993	5.930	1.3491	4.220
Perm. income: Mean	0.5083	2.624	1.0953	2.663
Std. deviation			0.2369	4.275
Homeownership in rural area (HR)				
Constant	0.5004	0.161	-9.9321	-1.652
Household size	0.0896	1.509	0.1058	1.171
Head's age: Mean	0.0797	2.784	0.2376	3.379
Std. deviation			0.0552	3.694
Head's age2	-0.0006	-2.355	-0.0017	-2.891
Male head	1.1510	5.208	1.7366	4.331
Head's education2	-0.7387	-4.463	-0.6905	-2.790
Head's education3	-1.3388	-5.650	-1.7510	-4.355
Region1	-1.9421	-10.500	-2.8002	-7.090
Region2	-0.2562	-1.670	-0.8902	-2.862
Casado	0.2553	1.244	0.7096	2.143
Married	0.8862	5.634	1.4041	4.457
Perm. income	-0.1927	-0.842	0.3251	0.808
Price	-0.5715	-4.078	-1.0487	-4.684
	Log-Likelihood: -4067.134 Log-Likelihood (constants only): -4610.912		Log-Likelihood: -4062.500 Log-Likelihood (constants only): -4610.912	

TABLE 3. Estimates of Choleski matrix for the mixed logit model of housing tenure-urban location

	Permanent income in HMIU	Head's age in HHU	Head's age in HRU
Permanent income in HMIU	0.2369 (0.0554)		
Head's age in HHU	-0.0564 (0.0153)	0.0002 (0.0793)	
Head's age in HRU	-0.0552 (0.0149)	0.0024 (0.0602)	0.0006 (0.0531)

Note: Standard errors are in parentheses.

TABLE 4. Marginal effects for the multinomial logit model of housing tenure-urban location

variables	Homeownership			Renting
	HU	MIU	Rural	
Permanent income	0.0596	0.0719	-0.0906	-0.0409
Household size	0.0020	-0.0111	0.0113	-0.0021
Head’s age	-0.0002	0.0105	-0.0012	-0.0091
Head’s age2	0.0000	-0.0001	0.0000	0.0001
Male head	-0.0245	-0.0661	0.1270	-0.0365
Head’s education2	0.0882	0.0007	-0.0985	0.0096
Head’s education3	0.1458	-0.0777	-0.1224	0.0543
Region1	-0.0180	0.0455	-0.1397	0.1122
Region2	-0.0376	-0.1055	0.0698	0.0733
Savings	0.0143	0.0381	0.0245	-0.0769
Married	0.0175	0.0780	-0.0414	-0.0541
HU owning price	-0.0334	0.0253	0.0041	0.0040
MIU owning price	0.0253	-0.1262	0.0618	0.0391
Rural owning price	0.0041	0.0618	-0.0760	0.0101
Renting price	0.0040	0.0391	0.0101	-0.0531

Note: The marginal effects are computed by averaging the individual sample observations

TABLE 5. Marginal effects for the mixed logit model of housing tenure-urban location

variables	Homeownership			Renting
	HU	MIU	Rural	
Permanent income	0.8959	1.0447	-1.3291	-0.6114
Household size	0.0046	-0.0298	0.0329	-0.0077
Head’s age	0.0972	0.2386	0.1987	-0.5345
Head’s age2	0.0177	-0.1094	-0.1292	0.2209
Male head	-0.0206	-0.0630	0.1214	-0.0378
Head’s education2	0.0309	-0.0060	-0.0254	0.0005
Head’s education3	0.0703	-0.0560	-0.0173	0.0029
Region1	-0.0121	-0.0044	-0.0245	0.0410
Region2	-0.0114	-0.0388	0.0275	0.0226
Savings	0.0072	0.0047	0.0069	-0.0187
Married	0.0143	0.0541	-0.0333	-0.0351
HU owning price	-0.7859	0.5812	0.1254	0.0793
MIU owning price	0.5727	-2.3982	1.4047	0.4208
Rural owning price	0.1230	1.3998	-1.7009	0.1781
Renting price	0.0761	0.4073	0.1713	-0.6546

Note: The marginal effects are computed by averaging the individual sample observations

TABLE 6. OLS estimates of housing demand equations with sample selection bias correction

Variables	Homeownership						Renting	
	HU		MIU		Rural			
	Param.	t-stat.	Param.	t-stat.	Param.	t-stat.	Param.	t-stat.
Constant	0.0415	0.021	1.5210	2.124	-0.6669	-0.450	-2.5667	-1.112
Perm. income	0.1773	1.644	0.1323	5.700	0.2119	4.319	0.7487	5.776
Trans. income	0.1311	2.175	0.1205	4.020	0.0905	2.158	0.2841	2.874
Housing price	-0.2864	-2.460	-0.2457	-5.461	-0.2483	-2.461	-0.6793	-6.363
Head's educ2	0.0881	0.528	0.0275	0.719	0.0548	0.920	0.3989	3.289
Head's educ3	0.4267	1.525	0.1140	2.099	0.1811	1.646	0.2882	1.414
Married	0.0113	0.156	0.0394	1.756	0.1585	3.059	0.2420	2.277
Prov. Capital	0.1560	3.007	0.1443	10.587	0.1387	1.376	0.2561	3.082
λ_1	-0.2916	-1.795	0.6029	1.294	0.6453	0.715	-1.2672	-1.333
λ_2	-0.3303	-0.523	-0.0462	-1.117	-0.7990	-2.392	0.0016	1.894
λ_3	1.9270	4.270	0.7860	5.758	-0.0837	-1.776	-0.7685	-1.662
λ_4	1.2181	2.042	0.2249	0.904	-0.0757	-0.145	0.2219	4.605
Adjusted R^2	0.2035		0.1778		0.1323		0.2862	
Log-Likelihood	-143.74		-1030.28		-475.68		-643.21	
Observations.	323		2830		803		504	

Note: The dependent variable is the quantity of housing services consumed in logarithm. Permanent and transitory income and housing price are in logarithm terms

TABLE 7. OLS estimates of housing demand equations without sample selection bias correction

	Homeownership						Renting	
	HU		MIU		Rural			
Variables	**Param.**	**t-stat.**	**Param.**	**t-stat.**	**Param.**	**t-stat.**	**Param.**	**t-stat.**
Constant	-0.5832	-0.336	-2.0137	-3.665	-2.7925	-2.009	2.1860	1.104
Perm. income	0.2646	3.395	0.2005	10.025	0.2565	5.684	0.3995	3.300
Trans. income	0.1146	1.840	0.1205	6.599	0.0900	2.144	0.2928	2.853
Housing price	-0.2470	-2.489	-0.0796	-2.153	-0.0945	-0.998	-0.6872	-7.287
Head's educ2	-0.0693	-0.888	0.0768	4.733	0.1005	2.467	0.5526	5.521
Head's educ3	0.0639	0.777	0.1552	6.561	0.1601	2.338	0.6804	5.005
Married	0.0008	0.013	0.0162	0.889	0.1189	2.639	-0.0688	-0.767
Prov. Capital	0.1488	2.833	0.1272	9.465	0.1033	1.018	0.1355	1.604
Adjusted R^2	0.137		0.158		0.114		0.224	
Log-Likelihood	-158.68		-1065.83		-485.87		-666.23	
Observations.	323		2830		803		504	

Note: The dependent variable is the quantity of housing services consumed in logarithm. Permanent and transitory income and housing price are in logarithm terms

TABLE 8. Income and price elasticities estimates for housing demand equations corrected for sample selection bias

Variables	Mixed logit model correction				Multinomial logit model correction			
	HHU	HMIU	HRural	Renting	HHU	HMIU	HRural	Renting
Perm. income	0.2424	0.1944	0.3184	0.4841	0.2391	0.2144	0.2744	0.5144
Trans. income	0.1311	0.1205	0.0905	0.2841	0.1389	0.1224	0.0925	0.2765
Housing price	-0.1469	-0.3309	-0.1094	-0.8641	-0.2121	-0.3559	-0.1470	-0.9240
Adjusted R^2	0.203	0.177	0.132	0.286	0.193	0.178	0.129	0.276
Log-Likelihood	-143.74	-1030.2	-475.68	-643.21	-145.73	-1028.4	-477.06	-646.43

TABLE 9. Housing demand estimates for owners and renters corrected by Heckman two step method and uncorrected

	Corrected for sample selection byHeckman's method				OLS			
	Owners		Renters		Owners		Renters	
Variables	Pa-	t-stat.	Pa-	t-stat.	Pa-	t-stat.	Pa-	t-stat.
Constant	-3.0092	-5.507	0.3028	0.138	-3.1030	-5.820	2.1860	1.104
Perm. income	0.2308	11.269	0.8294	5.341	0.2241	12.019	0.3996	3.300
Trans. income	0.1437	8.559	0.2714	2.726	0.1436	8.545	0.2928	2.853
Rel. Housing price	-0.0438	-1.050	-0.9013	-8.301	-0.0280	-0.765	-0.6872	-7.287
Head's educ2	0.0916	5.553	0.2923	2.427	0.0960	6.185	0.5526	5.521
Head's educ3	0.2176	9.236	0.2208	1.283	0.2247	10.331	0.6804	5.005
Married	0.0398	1.891	0.3375	2.691	0.0304	1.753	-0.0688	-0.767
Prov. capital	0.1909	14.634	0.2245	2.724	0.1903	14.597	0.1355	1.604
λ_T	0.0638	0.790	1.1045	5.427				
Adjusted R^2	0.227		0.279		0.227		0.224	
Log-Likelihood.	-1883.03		-642.51		-1887.85		-666.23	
Observations	3956		504		3956		504	

TABLE 10. Income and price elasticities estimates for owners and renters (housing demand equations corrected by Heckman's method and uncorrected)

	Corrected for sample selection by Heckman's method		OLS	
Variables	**Owners**	**Renters**	**Owners**	**Renters**
Permanent income	0.2271	-0.4241	0.2241	0.3996
Transitory income	0.1437	0.2714	0.1436	0.2928
Housing price	-0.0431	-1.1589	-0.0280	-0.6872
Adjusted R^2	0.227	0.279	0.227	0.224
Log-Likelihood	-1883.03	-642.51	-1887.85	-666.23

TABLE A1. Permanent income regression

Variables	Coefficients	t-ratios	Means
Constant	13.002	91.940	
Head's age	0.022	7.594	53.819
Head's age2	-0.0001	-7.083	3107.530
Head's education2	0.116	8.400	0.301
Head's education3	0.312	18.448	0.175
Savings	0.126	10.882	0.372
Indef. employment	0.181	12.344	0.780
Public mutuality	0.101	5.422	0.068
Private mutuality	0.109	5.620	0.077
Qualification	0.146	9.303	0.819
Members occupied	0.260	33.332	1.116
Capital or property rents	0.112	4.828	0.502
Secondary housing	0.130	8.296	0.143
Main income source1	0.164	1.401	0.607
Main income source2	0.117	1.002	0.387
Main income source3	0.502	3.117	0.002
Capacity to end month	0.226	19.386	0.461
Observations: 5259		**Adjusted R^2: 0.561**	
F: 421.65 (significance.: 0.000)		**Durbin-Watson: 1.779**	

Note: The dependent variable is the logarithm of current annual disposable income (LnY_c). The logarithm of permanent income (LnY_p) is the predicted value of this regression. The residual part is treated as the logarithm of transitory income (LnY_T). The explanatory variables included are the following:

Constant: constant term.

Head's age, Head's age2: Household head's age and its square value.

Members occupied: Number of household members occupied in the last quarter.

The next are dummy variables:

Head's education2 and 3: The highest education level attained by the household`s head is high school education, and college-university education, respectively.

Savings: Household possibility of saving monthly.

Indef. employment: Employment contract with indefinite duration.

Qualification: The household's head has some qualification for work (following the National Classification of Occupations: CNO94). The reference value (0) corresponds to the 9 group: workers without qualification.

Public and Private mutuality: Household's head is covered by a public or a private social mutuality, respectively.

Secondary housing: The household possesses secondary housing.

Capital or RE rents: The household obtains income from capital assets or real estate rents besides the main source of household's income.

Main income source1, 2, and 3: The main source of household's income is from working, from benefit pensions or subsidises, or from capital or real estate assets, respectively.

Capacity to end month: The monthly income flow of the household is enough to conclude month without economic problems.

TABLE A2. Hedonic price indices by regions within Spain (in logarithm)

Regions	Owning			Renting
	High level urban	Med.-Inf. urban	Rural	
Andalucía	13.447	13.095	13.216	12.552
Aragón / Rioja / Navarra	13.505	13.179	13.042	12.675
Asturias	13.392	13.107	13.256	12.092
Baleares	13.256	13.256	13.256	12.850
Canarias	14.177	13.420	13.380	12.355
Cantabria	13.673	13.355	13.264	13.259
Castilla y León	13.248	13.097	13.003	12.401
C. La Mancha / Extremadura	13.415	13.058	12.836	11.968
Cataluña	13.442	13.369	13.290	13.181
C. Valenciana / Murcia	13.265	12.973	12.962	12.380
Galicia	12.936	13.155	12.936	13.041
Madrid	13.776	13.573	13.776	13.382
País Vasco	13.219	13.219	13.219	13.647

TABLE A3. Descriptive statistics of housing characteristics

Housing characteristics	Mean	Standard deviation
Building type = 0	0.34	0.48
Age = 30 years	30	34.19
Age2 = 900	2073.85	43361.20
Rooms = 5	5.13	1.32
Rooms2 = 25	28.11	17.72
Area = 95	95.96	39.35
High urban = 0,1	0.083	0.28
Med-Inf. urban=0,1	0.67	0.47
Rural = 0,1	0.24	0.43
Provcap= 0	0.41	0.49
Dens1 = 1	0.51	0.50
Dens2 = 0	0.18	0.38
Dens3 = 0	0.32	0.47
Heating = 0	0.43	0.50

Note: The variables considered in the hedonic housing price equations (A1) and (A2) are::

Building type: A dummy variable indicating if the building structure type is single-family (value 1) or multi-family (value 0).

Age and *Age2*: Dwelling age and its square value.

Rooms and *Rooms2*: Total number of rooms including storages rooms, basements or lofts, and its square value.

Area: Lot size in square meters.

Heating: A dummy variable indicating if there is a heating system.

Provcap: A dummy variable indicating if it is located in an area corresponding to a Spanish province.

Dens: This variable reflects population density with three dummies:

> *Dens1*: It indicates a densely populated area. It corresponds to an area comprising diverse adjacent towns, having each one a population density over 500 inhabitants by square kilometre, and a total population over 50.000 inhabitants
>
> *Dens2*: It indicates an area with a medium population density. It comprises diverse adjacent towns that does not belong to a densely populated area, having each one a population density over 100 inhabitants by square kilometre and verifiying one of two conditions: a total population over 50.000 inhabitants, or they are situated nearby a densely populated area.
>
> *Dens3*: It collects those dwellings not included in Dens1 or Dens2.

Degree of urbanization: This variable reflects the urban location alternatives considered in this paper: h*igh level urban medium-inferior level urban and, rural location.*