A comparative analysis of different models of the tenure choice of a house

Josep Maria Raya
Escola Universitaria del Maresme and Department of Economics and Business
Universitat Pompeu Fabra
C/Ernest Lluc s/n 08302, Mataró (Barcelona)
Tel:931691611
josep.raya@upf.edu

Jaume Garcia
Department of Economics and Business
Universitat Pompeu Fabra
C/Ramon Trias Fargas 25, 08005 Barcelona (Barcelona)
Tel:935421748
jaume.garcia@upf.edu

Abstract: The individual's choice of house tenure is closely bound to other vital individual decisions of great importance for the subject and for the country in which he resides. The present paper includes a comparison of the different approaches that have been used in the economic literature in the analysis of the individual's tenure choice of housing. This approach gives rise to groups according to their dependent variable, that is: the classic models, the Markov chain models, the "recent-movers" models, the sequential models (those whose dependent variable is the tenure), the duration models and the "choice-based" models (those in which the dependent variable is the transition from renting to the position of ownership). The results show the goodness of fit of the models in which the variable to be explained is the transition from renting to the position of ownership.

JEL classification: R21, C25, C41, C52, C53
Key words: tenure choice, comparative analysis, survival models, classic models.

Comparative analysis of different housing tenure models

Introduction

An individual's housing tenure choice is closely bound to other important life-cycle decisions of the person concerned. Furthermore, the percentage of home owners in a country has consequences for several economic and social aspects of that country. Let us consider some examples to illustrate this statement. First, the combination of a culture of home ownership, a small rental market and high housing prices leads to a delay in young people leaving home in Spanish society. In addition, this phenomenon can affect a society's patterns of marital status and birth rate. In this respect, Haurin et al (2002), for example, holds that ownership is associated with several positive socioeconomic impacts. In this sense, and also for homeownership, DiPasquale and Glaeser (1999) show greater social commitment and more care about their living environment and Green and White, (1997), show a positive impact on the education of children. Dietz and Haurin (2003) review the literature on the social effects of home ownership. Lastly, ownership entails higher mobility costs and therefore less scope for geographical mobility (Shelton, 1968; Oswald, 1997 and Ortalo-Magné and Rady, 2002). As a result, a country with a higher percentage of home ownership runs a higher risk of structural unemployment and hence lower potential growth and financially weaker households in the face of interest rate rises.

In view of the above, it is of interest to know the tenure choice of individuals and its determinants. The models used in the economics literature are diverse, and range from the classical binary choice models to duration models. The purpose of the present paper is to estimate, using the same data for the same country, Spain, the different models that have been used in the housing economics literature to describe individuals' housing tenure choice. Thus, the paper serves firstly as a review of several different models of housing tenure choice. Furthermore, the comparative study of the results of the estimates of these models will illustrate the differences in results that occur depending on the specification used. As far as we know, this is the first time this type of exercise has been done in the economics literature.

The paper is structured as follows. In section 1 we present a detailed analysis of the economics literature on housing tenure choice and its evolution, with special attention to the issue of

econometrics. We then go on to present the empirical model and the data on which the comparative study will be based. In section 3 we will present the results of the different estimates, and lastly we will conduct a comparative study of the results of the various models and finish with some final conclusions.

1. Alternative approaches to modelling tenure

The first tenure models follow the line initiated by King (1980), who uses an empirical model to model housing tenure and the demand for housing services jointly, as they have their origin in the same ranking of preferences. Henderson and Ioannides (1985) constitutes a first attempt to provide these types of models with a rather more sophisticated theoretical foundation. The intertemporal dimension, to be taken into account in any life-cycle decision, is incorporated in Zorn (1988), Goodman (1995), Ioannides and Kan (1996). These papers complete Henderson and Ioannides' (1985) dynamic programming problem and enable us to obtain the tenure choice made by each individual in each period, and not only in a particular period, always within a model that aims to account for the housing services demand made by the consumer.

Econometric problem

Turning our attention to the econometric problem, chronologically one of the first studies to address these types of models is King (1980),. King (1980) is based on a vector of housing services consumption according to the tenure status, v , that allows *J* types of tenure:

$$v = (x_1 y_1, ... x_J y_J) \tag{1}$$

where x_J is the amount of housing services consumed in status *j*, and y_J takes the value 0 or 1 depending on the tenure choice, with $\sum_{j=1}^{J} y_j = 1$. The probability of observing household *i* with tenure status j is denoted by p_{ij} and the density function of the housing services demand conditioned to the tenure choice by $f_{ij}()$. The log-likelihood function for any random sample is given by:

$$L(v) = \sum_i \sum_j y_{ij} \log p_{ij} + \sum_i \sum_j y_{ij} \log f_{ij}(x_{ij}) \quad (2)$$

In order to perform the estimation we need the choice of tenure and service consumption y_j, x_j and a theory about the determinants of p_{ij} and $f_{ij}()$. These types of tenure models, including those presented in Henderson and Ioannides (1985), Goodman (1988), Bourassa (1995), are estimated using discrete choice models. Thus, by using all the individuals in a cross-sectional sample, they are capturing consumption behaviours on the basis of past (unobservable) prices, costs and incomes. That is to say, the individuals are being observed at a distance from the time at which the decision concerned was made.

Thus, the econometric estimates of these tenure models, henceforth referred to as *classical tenure models*, can be considered biased (Painter, 2000): instead of capturing the influence of the explanatory variables on the tenure choice they actually describe the profile of the owner or tenant. In this respect, these models, which are a good choice if the aim is to describe tenure status of the population as a whole, may largely reflect the lagged effects of past choices. This is the biased source, since in cross-sectional data, current data is not likely to reflect past choices among homeowners. This is especially so if no control is introduced to account for the previous tenure status. Previous tenure has an obvious impact on mobility and current tenure. Therefore, previous tenure status is a good indicator of preferences with regard to tenure, and also financing and other aspects. Individuals are more likely to be owners in the current period if they were in the previous period, among other reasons because it shows their preference for ownership, because of the high adjustment costs for home owners, and because, *ceteris paribus*, they are more likely to obtain funding. Kan (2000) highlights these aspects in his multi-equation model of tenure and mobility. Thus, if we have longitudinal data for a set of households, we can model housing tenure using a first-order Markov chain[1].

A first-order Markov chain is characterized completely if we specify the probability of transition defined by the probability of individual *i* being in state *k* at time *t* if he or she was in state *j* at time *t-1*

[1] Higher-order autoregressive schemes not only have a considerable computational cost in short longitudinal series such as the existing housing data but furthermore are not justified by the mainstream economics literature.

($P_{jk}^{i}(t)$).[2] Using panel data and following Jiménez et al (1996) for an explanatory model of the decision to strike, the model to estimate with the presence of the lagged endogenous variable for the latent criterion (y*) of tenure choice (y) is as follows:

$$y_{it} = \gamma y_{it-1} + \alpha x_{it} + f_i + \varepsilon_{it} \tag{3}$$

$$y_{it} = 1 \ \ (y_{it}^{*} \geq 0) \tag{4}$$

where x is the vector of determinants influencing the latent criterion, ε is the error term, α and β are the parameter vectors and f_i is an individual effect. If the dynamics and the individual effects are relevant factors in explaining tenure choice, the estimators obtained in a logit model are inconsistent, due to the problem of correlation between the lagged endogenous variable and the unobservable individual effect. In this model, the lagged endogenous variable cannot simply be considered a regressor (Maddala, 1987). Under these premises, the simplest alternative to apply to panel data (other less optimal alternatives for this case include the logit model with specific unobservable effects and the probit random effects model) is the linear probability model. Although this alternative has some well-known shortcomings, it allows us to introduce the lagged endogenous variable as an explanatory factor and use instrumental variable methods to obtain consistent estimators (Arellano and Honoré, 2001). Additionally, we can control the specific effects, taking first differences in (3):

$$\Delta y_{it} = \gamma \Delta y_{it-1} + \alpha \Delta x_{it} + \Delta \varepsilon_{it} \tag{5}$$

The parameters of (5) can be estimated using the generalized method of moments (GMM, Arellano and Bond, 1991). This procedure controls for the fact that $\Delta\varepsilon_{it}$ is heteroskedastic and serially correlated. Furthermore, empirical analysis allows us to test the importance of the specific effects (Sargan statistic difference test, Arellano, 1993). This test is very useful when the vector of explanatory variables includes the lagged endogenous variable. By construction, this variable is correlated with the individual effects, and so in principle the levels model does not yield consistent

[2] See Amemiya (1985) for a description of these models.

estimators. However, under the null hypothesis that these effects are irrelevant, both models yield consistent estimators. Therefore, the Sargan statistic is tested under the null hypothesis that both estimates are consistent and the alternative that only those obtained in the first-differences model are. The difference between the two statistics is distributed as a χ_r^2, where *r* is the number of additional orthogonality restrictions of the levels model in relation to the first differences model.

Nevertheless, whether or not we control for the previous tenure status, we are still observing individuals at a distance from the time of the decision we are interested in. In contrast, the market is observed at the right moment for individuals who have just moved house. Henderson and Ioannides (1985), who work from King's joint modelling of tenure and consumption, distinguish two types of specifications. The first of these is similar to King's, but uses a sample of recent movers (households who have recently moved house) rather than a sample of all the households in the survey. Therefore, the study focuses solely on the tenure decision of those individuals who are behaving in response to current income, transition costs and prices.

The problem now is one of sample selection bias, as in this case the tenure choice is only observed if the household moves. This is addressed by Painter (2000), Haurin et al (2002), Painter et al (2004) and Painter and Gabriel (2006). Painter (2000) works from a latent variable *y** that is regressed on a vector of demographic and economic variables.

$$y^* = \beta X_i + \varepsilon_{1i} \tag{6}$$

However, we only observe the binary output, *y*=1 if *y**>0 and *y*=0 otherwise. Furthermore, we only observe *y* if the individual has moved, i.e., *m*=1, where *m** comes from the ratio below:

$$m^* = \gamma z_i + \in_{2i} \tag{7}$$

where *m*=1 if *m**>0 and 0 otherwise, and z_i are demographic and economic variables that do not necessarily coincide with X_i. It is assumed that ε_{1i} and ε_{2i} are distributed jointly with correlation coefficient ρ. All the above allows a maximum likelihood estimate of:

$$\sum_{i \in S}^{vi=1} \ln\left[\Phi_2(x_i\beta, z_i\gamma, \rho)\right] + \sum_{i \in B}^{vi=0} \ln\left[\Phi_2(-x_i\beta, z_i\gamma, \rho)\right] + \sum_{i \in B} \ln\left[1 - \Phi_1(z_i\gamma)\right] \tag{8}$$

where S is the set of recent movers and subscripts 1 and 2 denote the cumulative and bivariate normal distribution function respectively.

Goodman (2003), which monitors a panel of households that stay in their home, was the first attempt to study empirically tenure choice as a life-cycle dicison, at least as regards its date of formulation, albeit not that of publication. The estimation method is considered an econometric approach to the multi-period model derived by the same author, Goodman (1995), which shows the equilibrium conditions. Period by period, individuals compare the cost of staying in their home (in terms of the utility they lose in relation to the home they want) with the costs of moving. Analogously, the costs of being a home owner are compared with those of renting. The final objective is to obtain the impact on the elasticities of both those who stay in their home during the period concerned (stayers) and those who do not (movers). The literature until then had only obtained partial elasticities, of either stayers (classical models) or movers (recent-movers models). Goodman (2003) provides a theoretical and empirical framework to study this phenomenon by estimating sequential bivariate probit models.

However, the empirical model presented in Goodman (2003) only captures the tenure decision at the moment it is made for those individuals who move, i.e., those who leave the sample (movers). The stayers made their tenure choice in some previous period. So, in order to model an individual's tenure status at each time *t*, or to be more exact, the probability of the individual retaining a particular tenure status in which he or she has been for *t* periods, the most suitable method is to use discrete time survival or duration models. In these types of models, time is measured in discrete intervals. That is to say, the state of an individual *i* is observed from year $t=1$ to year $t=T_i$. For the case that concerns us here, we observe whether the individual owns or rents his or her home from year $t=1$ to year $t=T_i$. In T_i, unless we have ceased to observe the individual at some stage, the event is complete and we have obtained a sequence of states for that period. The time of survival in a particular state (ownership or tenancy) T_i is a discrete random variable. The hazard function that comprises the probability of leaving a state in period *t* conditioned by having been in that state (h_{it}) for *t* periods is:

$$h_{it} = \Pr(T_i = t \mid T_i \geq t) \tag{9}$$

The likelihood function for the entire sample, once a new binary variable $y_{it}=1$ is defined if individual *i* reaches the situation of success (completes the event) in year *t* and $y_{it}=0$ otherwise, can be written as:

$$\log L = \sum_{i=1}^{n}\sum_{j=1}^{t_i} y_{ij}\log(h_{ij}/1-h_{ij}) + \sum_{i=1}^{n}\sum_{j=1}^{t_i}\log(1-h_{ij}) \\ -\sum_{i=1}^{n}\sum_{j=1}^{t_i}\left[y_{ij}\log h_{ij} + (1-y_{ij})\log(1-h_{ij})\right] \tag{10}$$

The above expression has exactly the same form as the standard likelihood function for a binary choice model in which y_{it} is the dependent variable and the data have been reorganized; specifically, from having one data per individual to having one data per year in which the individual can make a transition (Jenkins, 1995). The principal duration models estimated for housing are: Henderson and Ioannides (1989), Di Savio and Ermisch (1997) and Ahn (2001), Smits and Mulder (2008).

Lastly, in view of the little weight given to transitions in a person-period sample, Zorn (1988) suggests using a choice-based sampling technique (based on the endogenous variable), thus giving them a greater presence in the sample. According to Coslett (1978), if we use a logit model, due to its functional form, non-random smapling only influence the constant term, which furthermore is straightforward to estimate consistently.

2. Empirical specifications for each model

In the present paper we will make estimates of six procedures used in the literature to analyse the tenure decision made by individuals. Specifically, the list of models to be estimated is: the classical tenure model (*classical model*), which follows from (2) and describes the probability of an individual being a home owner at a given time depending on his or her sociodemographic characteristics; the model that controls for previous tenure, based on a first-order Markov chain with individual effects, the specification for which comes from (5) (*Markov model*); the model estimated only for those individuals who are in equilibrium, i.e., who have recently moved and have therefore made their tenure choice, which follows from (8) (*recent-movers model*); the sequential bivariate probit model, which estimates tenure choice and mover-stayer status jointly (*sequential model*); the duration model, which follows from (10) and reflects the decision to make a transition from one tenure status to another at a particular moment in time, conditioned to not having been made before (*duration model*);

and lastly, the transition model, which uses a sampling based on tenure choice, giving it greater weight in the sample (*choice-based model*).

The data used are taken from the European Community Household Panel (ECHP), in particular, waves from 1994 to 1999. ECPH is a standardized survey that is carried out in the European Union. The database contains information on housing and socioeconomic characteristics of the household and its members. A general overview of the ECHP is given in Peracchi (2002).

Focusing on the types of models that we wish to analyse below, the Household Panel presents no problem at all for preparing a classical tenure model, a Markov model, a recent-movers model or a sequential model. However, six periods would appear to be insufficient to prepare a duration model. Yet one of the components of the ECHP is key for obtaining retrospective information: "year since which respondent has lived in present home". This question enables us to extend the duration model. In this way, by simply introducing the inevitable assumption that during the "extra-sample" time in which the respondent has lived in the building he or she has not changed tenure status, we can obtain the time the individual has remained in a given state (home ownership or tenancy) until the time of transition, if applicable. By analogy, once the duration model is extended, the estimation of a choice-based model will be fully feasible.

A general sample was compiled with Spanish households that stayed in the panel for more than one period, some households being removed from the sample if information was missing for some of the more important variables..

The sample consists of 6778 households, almost 82% (87% if we only consider ownership and tenancy status) of which were home owners. Furthermore, 40.36% made at least one move over the five periods observed, regardless of whether they retained the same tenure status or made a transition, and they were observed on average for 3.64 periods. Lastly, the transition which we will study, the transition from tenancy to home ownership, was made by 5.85% of the households during the period.

Given different objectives and definition of dependent variables, we categorised the six models into two groups: models in which the dependent variable is the decision between home ownership and renting (classical, Markov, recent-movers and sequential) and models in which the dependent variable is the transition made by the individual at a given moment from renting to home ownership (duration and choice-based models). Below we define the variables that form part of the various models.

The definition of the dependent variable for the main equation of all the models in which the dependent variable is the decision between home ownership and renting is as follows:

Tenure status: A variable that takes the value 1 if the individual is the owner and 0 otherwise. In the case of the classical model the tenure status is taken as that observed at the time of the cross section, 1995, whereas in the case of the recent-movers model the tenure status is taken as that observed after the move.

Mover: For the slection equation of the recent-movers model. This takes the value 1 if the individual has made some move or change of residence during the period of up to five years in which he or she has been in the sample, and 0 otherwise. These moves may or may not give rise to changes in tenure status.

Stayer: For the second equation of the sequential model. Takes the value 1 if the individual stays in the same home in the period following that in which the tenure status is observed, and 0 otherwise.

Transition: A variable that takes the value 1 for those individuals who make the transition from tenancy to home ownership, and only at the time this transition occurs. It takes the value 0 in all other cases.

In turn, the explanatory variables included in the various models estimated are as follows:

Tenure status in previous period: This is the autoregressive component that gives rise to the first-order Markov chain.

Income: Following Goodman (1988), we estimate the permanent income and transitory income of the household on the basis of the current monetary income of the head of household. We use this approach because including actual income as a determinant of the tenure choice causes endogeneity. Moreover, according to the permanent income hypothesis, decisions on durable goods consumption are likely to be a function of permanent income instead of current income. Thus, assuming home ownership as a long-term decision, home ownership is correlated with life cycle characteristics such as permanent income. In accordance with Goodman's approach, permanent income is the predicted value of a regression of household income on a set of demographic and human capital characteristics. Transitory income is calculated as the residual of observed household income and predicted income. The variable is expressed in thousands of Euros

Gender: Dichotomous variable that takes the value 1 if the main earner is a man and 0 if the main earner is a woman.

Age: Age in years of the main earner. Quadratic profiles are usually specified.

Household size: Number of people who consume in the household. Following Kan (2000), in order to capture whether a transition in *t* occurs in response to a recent change in some characteristic of the household, we also specify this variable as the change in the number of members (if it is positive) undergone by the household in the five years previous to the transition. This definition will be used in the duration and choice-based models.

Marital status: Five dummy variables are defined for the five states established in the ECHP for the main earner: married, separated, divorced, widowed or single. Similarly to the above case, this is specified as a dichotomous variable taking the value 1 if the individual has changed marital status in the last five years and 0 otherwise.

Employment status: Specified in the form of five dummy variables corresponding to the five categories of this variable in the survey. These are, in order: working 15 hours or more per week, working less than 15 hours per week, unemployed, long-term unemployed, and out of the labour force. Along similar lines to the previous cases, again this variable is defined as a dummy variable with the value 1 if the individual changed from being in one of the last three of the above categories to being in one of the first two during the period prior to the transition, and 0 otherwise.

Education: Highest level of education completed by the individual. It is presented with a single specification in three dummy variables defined through the three categories of educational level considered in the panel: university or equivalent, secondary education, and primary education or lower.

Relative price: This is the difference in euros between yearly payments for ownership (in terms of mortgage amortization) and tenancy in the period prior to the transition or move. Note that for each individual we only observe one of these two prices: amortization for home owners and rent for tenants, both flow variables. As a result, in order to estimate the other price we use Heckman's two-step procedure. The first step consists in estimating a selection equation in which the dependent variable (tenure status) depends on a set of socioeconomic characteristics of the individual. In the second step, once selection is accounted for, and therefore only for home owners in order to obtain the rent and only for tenants in order to obtain the amortization, we regress the price on a set of characteristics of the home that are present in the ECHP. These characteristics are: number of bedrooms, type of dwelling, age, presence of a series of features such as a terrace, bathroom or

heating, and the household's perception of a series of aspects of the area, such as crime, noise or pollution, measured in dummy variables. We also control for the type of landlord. This second equation is therefore a conventional hedonic regression (results can be obtained by author's request).

Duration: For those households that made a transition, this variable takes the value given by the difference, in terms of numbers of years, between the moment in time at which the observation registered the transition of observed tenure and the moment at which the individual began to live in the rented home. However, for those households that were not observed to have made any transition, the variable *duration* takes the value of the difference, again in terms of number of years, between the moment at which the individual began to live in the home and the moment at which the individual was observed for the last time. This variable in itself constitutes an approximation of the transaction costs, among other aspects. It is assumed that the longer a household has been living in a home, the more expensive it is to leave it.

Number of bedrooms: This variable is intended as an approximation of the transaction costs (Zorn, 1988). The number of bedrooms in the previous home is intended as an approximation of its size.

The descriptive statistics of all these variables are presented in Table 1. The mean and the standard deviation of each variable are differentiated for the individual and for person-period data.

[TABLE 1]

3. Results

We will now go on to compare the various models estimated.. In terms of their objectives and the definition of the dependent variable, the classical, Markov chain and recent-movers models and Goodman's bivariate probit approach are all models that deal with the issue of being a home owner in relation to renting. Therefore, they explain the characteristics that describe the home owner. In contrast, both the duration model and the choice-based model focus on the transition made by the individual at a given moment from renting to home ownership. Therefore, these models explain the reasons for this change. We understand that the latter case is more interesting, especially because the literature on tenure always states that its objective is to study the decision between home ownership and renting, and it is in the latter models that this decision is captured most clearly at the moment of transition. As we pointed out previously, we will make homogeneous comparisons, i.e., comparisons between the models within each of these main groups. In particular, we will make an intuitive

comparison of the estimates on the basis of the significance of their coefficients and then a more robust comparison in terms of predictive ability and marginal effects. Only at the end will we offer an intuitive vision of the perceived differences between the two main groups.

Table 2 presents a summary of the significances of the variables in each model. They are reported as "S" if the variable is significant, "I" if the variable is not significant, and "-" if the variable is not included in the model.

In terms of the significance of the coefficients (Table 2), permanent income, age, marital status and primary education in relation to university education are significant in all the models in which the dependent variable is home ownership. Relative price is not significant in the main equation of the recent-movers model (but it is in the selection equation). The explanation for this would appear to be that the price of ownership with respect to tenancy only affects tenure choice indirectly, through mobility. Employment status only appears as significant in the recent-movers model, whereas household size only appears as significant in the classical model and the recent-movers model. The effect of the variable *previous tenure status* appears as significant in the model that incorporates it. To sum up, the significance pattern seems to be marked by the sequential model based on Goodman's approach.

Regarding the significance of the second group of models, in which the dependent variable is the transition from the status of tenancy to that of home ownership, practically the same variables appear as significant in the duration model and the choice-based model. These variables are: income, age, marital status, household size, relative price, duration and number of bedrooms.

[TABLE 2]

Going on to consider the predictive ability of the models whose dependent variable is tenure status, it should be noted that the percentage of home owners is the same in all cases, as we use the same sample. We observe in Table 3a that, for a threshold of 0.5 (i.e., assigning a value of 1 when the prediction is greater than 0.5 and 0 otherwise), the classical model, the recent-movers model and the sequential model were accurate between 80% and 88% of the time. However, they show a percentage of correct predictions of less than 15% (4% for the recent-movers model and 5% for the classical model) in the case of individuals whose tenure choice is to rent. On the other hand, the Markov model shows the highest total percentage of correct predictions (96%), with a considerable increase in the

accuracy for renting (77%). Lastly, the percentage of correct predictions for home owners is high in all the models, albeit slightly higher in the Markov model.

As regards those models in which the dependent variable is the transition from tenancy to home ownership in a particular period, the aim will be to predict the probability of the individual making this transition in one of the periods observed. Thus, with a threshold of 0.05, as observed in Table 3b, the predictive ability of the duration model is lower (61.5%) than that of the choice-based model (68.3%). However, the duration model predicts both groups reasonably well. It provides an accurate prediction in 62.5% of cases in which the households make no transition and in 55.5% of cases in which the households do make the transition. For its part, the choice-based model is accurate in 68.5% of cases in which the households make no transition and in 65.1% of households that do make the transition.

[TABLE 3a] [TABLE 3b]

Therefore, in terms of predictive ability, we conclude that among the models in which the dependent variable is tenure status, the Markov model stands out for its predictive ability. However, when the objective is to predict whether the household will make a transition from the status of tenancy to that of home ownership at a particular moment in time, which would appear to be more complicated from an intuitive point of view, both models show reasonably high success rates. In this they rank only slightly lower than the Markov model, and furthermore strike an even balance between households that do not decide to make a transition and those that do.

Lastly, we will make a comparison of the results of the models as a function of their marginal effects. First of all, we have to clarify that our paper considers objectives and predictive capacity as the main reasons to choose between tenure choice models. In this sense, marginal effects are used to reinforce objectives and predictive capacity criterion rather than constituting a criterion in itself.

We understand by marginal effect the change in the expected value of the dependent variable, in this case a probability, on making a marginal change in the explanatory variable. Or analogously, the derivative of *y* (dependent variable) as a function of *x* (explanatory variable), (dy/dx). In discrete choice models, the marginal effects are not constant and depend on the point of the explanatory variable at which they are evaluated. In this case, they will be evaluated at the mean of the explanatory variables, except for the dichotomous variables whose marginal effects will represent the change in the dependent variable when the explanatory variable changes from 0 to 1. Tables 4a and

4b show a comparison of the marginal effect of each independent variable measured at its mean value. To the right of each marginal effect we see the percentage probability of its endogenous variable, tenure status, taking the value 1, measured at the mean value of the exogenous variables. Our comments will refer to these values, and only for those variables whose coefficients were statistically significant at 5% (Table 2).

With regard to the models in which the dependent variable is tenure status, marital status stands out as the variable whose change has the greatest effect on tenure status (Table 4a). The categories that do most to reduce the probability of being a home owner are those of being separated and being divorced, with 8% and 18% in the case of the classical and recent-movers models, 24% and 27% in that of the sequential model, and 38% and 10% in that of the Markov model. In this last model, the effect of having been a home owner in the previous period, with respect to not having been one, is notable, causing a 78% increase in the probability of being a home owner in the current period. The effect of a marginal change in the permanent income ranges from 1.43% in the classical model and the Markov model to 4.67% in the recent-movers model. Finally, the effect of a marginal change in the relative price lies around 1%, and a change from having higher education to having only primary education increases the probability of being a home owner by between 0.5% (Markov model) and almost 7% (recent-movers model).

Comparing the marginal effects of the different models, in general we observe that they are slightly higher for most variables in the recent-movers model and the sequential model, and more concentrated in just a few variables in the Markov model. The presence of a particular tenure status in the previous period seems to be associated with a slight reduction in the marginal effects of the rest of the variables, with the exception of marital status and employment status. This is not surprising, as the same variables that explain current tenure probably also explained previous tenure. Thus, the lagged tenure status tends to overwhelm the marginal effects of the rest of variables. This result lead us to conclude that although the higher predictive ability of the Markov model, doesn't help us to understand the determinants of the tenure choice.

On the subject of marginal effects in the models in which the dependent variable is the transition from tenant to owner status in a particular period (Table 4b), first of all it should be noted that, in terms of the mean value of the probability of being a home owner, they are higher than in the previous models. To our understanding, this is an indication that these models truly capture the

moment of choice and the effect of the different variables that influence it. Of all these effects, that of marital status stands out as that of greatest importance (from 73% to 127%), unless among the categorical variables. This shows the temporal coincidence of the two decisions within the life cycle of the individual. The high correlation of the two decisions with the fact of leaving the parental home justifies a result already expressed by Ahn (2001). This is no longer the case when the study is conducted in the field of second homes (Garcia and Raya, 2006). Dieleman and Everaers (1994) found that the transition to homeownership mainly takes place when stability in both income and household situation has been reached. Of the remaining marginal effects of the significant variables, the most outstanding are the effect of income, with a marginal effect of between 17% and 24%, and that of household size, between 21% and 40%. On the other hand, the marginal effects in the duration and choice-based models are very similar, although slightly higher in almost all cases in the duration model.

[TABLE 4a] [TABLE 4b]

In the light of the above tables, it is straightforward to find greater significances and marginal effects in the models in which the dependent variable is the transition from tenure in the form of renting to home ownership.

4. Conclusions

In the present paper we have attempted to conduct a review and a comparison of the different approaches that have been made in the economics literature to the analysis of individuals' tenure decision regarding their main home. These approaches can be grouped depending on their dependent variable and analysed on the basis of their six main formats. The first group, in which the dependent variable is tenure status, consists of the classical models, the Markov model, the recent-movers models, and the sequential model. The second group, in which the dependent variable is the transition from the tenure status of renting to home ownership, consists of duration models and choice-based models.

To sum up, in connection with the models in which the dependent variable is tenure status at a given moment in time, no major differences are detected regarding the significance of the variables. However, with regard to their marginal effects, the recent-movers model stands out. In this model, the greater marginal effects of income, employment status and education are evident. The explanation for

this is that this model captures tenure status for those individuals who move at that moment. In this respect, to a large extent it captures the real effect of the explanatory variables at the moment the tenure decision is made. In turn, in the sequential model the marginal effect of relative prices is notable, and it shows higher marginal effects than the classical and Markov models in most of the other explanatory variables. This is the case because the sequential model also captures tenure status for those individuals who have left the sample (movers) and therefore are in the process of making a tenure choice. As regards results, the marginal effects are slightly lower than those observed in the recent-movers model. Lastly, the Markov model highlights the effects of previous tenure and household size. In the rest of the explanatory variables the marginal effects are lower than those observed in the recent-movers and sequential models. This is so because the Markov model does not capture the true tenure decision at any moment, although it takes into account the individual's history through previous tenure, which captures his or her tenure preferences. In all three of these models, marital status shows a considerable marginal effect, which does not occur in the classical model.

If we compare the results of the models in which the dependent variable is the fact of making a transition from renting to home ownership, we observe identical significances and very similar marginal effects. To mention some differences, in the duration model the marginal effects of income, household size and marital status are slightly higher, whereas in the choice-based model the marginal effects of age, education, number of bedrooms and duration are higher.

In general, the results provide evidence of the best performance of the models in which the variable to be explained is the transition from tenancy to home ownership. This best performance is observed not only in terms of objectives but also in terms of predictive ability and marginal effects more close to what we expect. These models seem to capture the decision of interest by following the whole process undergone by the individual, and not just one particular state (whether that is the moment of a transition or some other) without knowing its past. The models in which the dependent variable is tenure status at a particular moment in time do not explain the ownership decision and only describe the profile of the home owner at a certain moment in time, which may be close to or far from the moment of the actual decision.

In terms of predictive ability, when the objective is to predict the change of tenure status at a particular moment in time, which would appear to be more complicated from an intuitive point of view, both models (duration and choice-based models) show reasonably high success rates. In this

they rank only slightly lower than the Markov model and furthermore strike an even balance between households that do not decide to make a transition and those that do.

Furthermore, these types of models make it possible to use variables with high explanatory power such as duration in a particular tenure status before making a transition. In addition, as the decision of interest is appropriately captured, we observe an increase in the magnitudes of the marginal effects in comparison with their magnitude in the models in which the dependent variable was housing tenure status. This is the case of the notable increase in the effect of marital status, household size and income. In fact, in Goodman (2003) the owner and renter demand estimates verify criticisms versus cross-section methods, as well as single year income measures.

As we mentioned earlier, the aim of this study was to review and compare the various different models that have appeared in the economics literature to explain the determinants of tenure choice. As further research, one interesting exercise would be to analyse, using a single database such as the ECHP, the determinants of convergence or divergence at European level, both for the probability of being a home owner and for that of making a transition from renting to ownership at a particular time.

References

AHN, N. (2001): "Age at first-time homeownership in Spain", *Documento de trabajo FEDEA 2001-23*.

AMEMIYA, T. (1985): *Advanced econometrics*. Harvard University Press.

ARELLANO, M. (1993): "On the testing of correlated effects with panel data", *Journal of Econometrics 59*, 87-97.

ARELLANO, M. and BOND, S. (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies 58*, 277-297.

ARELLANO, M. y HONORÉ, B. (2001): "Panel Data Models: some recent developments", *Handbook of Econometrics 5, chapter 53.*

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

COSLETT, S. (1978): "Efficient estimation of discrete-choice models", in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications* (Cambridge, MIT Press), 51-111.

DI SALVIO, P. and ERMISCH, J. (1997): "Analysis of the dynamics of housing tenure choice in Britain", *Journal of Urban Economics 42*, 1-17.

DIELEMAN, F. M. and EVERAERS, P. C. J. (1994): "From renting to owning: life course and housing market circumstances", *Housing Studies 9*, 11–25.

DIETZ, R. D. and HAURIN, D. R. (2003): "The social and private micro-level consequences of homeownership", *Journal of Urban Economics* 54, 401–450.

DIPASQUALE, D. and GLAESER, E. L. (1999): "Incentives and social capital: are homeowners better citizens?" *Journal of Urban Economics* 45, 354–384.

GARCIA, J. and RAYA, J.M. (2006): "La fiscalidad sobre la renta y el régimen de tenencia de la vivienda. Un análisis empírico para el caso español", *Papeles de Economía Española*, 198-213.

GOODMAN, A.C. (1988): "An econometric model of housing price, permanent income, tenure choice and housing demand", *Journal of Urban Economics 23*, 327-353.

GOODMAN, A.C. (1995): "A dynamic equilibrium model of housing demand and mobility with transactions costs", *Journal of Housing Economics 4*, 307-327.

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

GREEN, R. and WHITE, M. (1997):"Measuring the benefits of homeownership: effects on children", *Journal of Urban Economics 41*, 441–461.

HAURIN, D.R., PARCEL, T. and HAURIN, R.J. (2002): "Does homeownership affect child outcomes?", *Real Estate Economics 30*, 635-666.

HECKMAN, J.J. (1979): "Sample selection bias as a specification error", *Econometrica 47*, 153-162.

HENDERSHOTT, P.H. and PRYCE, G., (2006): "The sensitivity of homeowner leverage to the deductability of home mortgage interest", *Journal of Urban Economics 60*, 50–68.

HENDERSON, J. and IOANNIDES, Y. (1985): "Tenure choice and the demand for housing", *Economica 53*, 231-246.

HENDERSON, J. and IOANNIDES, Y. (1989): "Dynamic aspects of consumer decisions in housing markets", *Journal of Urban Economics 26*, 212-230.

JENKINS, S. (1995): "Easy estimation methods for discrete time duration models", *Oxford Bulletin of Economics and Statistics 57*, 129-138.

JIMENEZ, S., LABEAGA, J.M. and MARCO, M. (1996): "Algunos factores explicativos de la existencia de huelgas durante la negociación colectiva en España", *Investigaciones Económicas 20*, 217-241.

KAN, K. (2000): "Dynamic modelling of housing tenure choice", *Journal of Urban Economics 48*, 46-69.

KING, M. (1980): "An econometric model of tenure choice and demand for housing as a joint decision", *Journal of Public Economics 14*, 137-159.

MADDALA, G.S. (1987): "Limited dependent variable models using panel data", *The Journal of Human Resources 22*, 307-338.

OSWALD, A. (1997: "Thoughts on NAIRU (correspondence)", *Journal of Economic Perspectives, 11*, 227–228

PAINTER, G. (2000): "Tenure choice with sample selection: differences among alternative samples", *Journal of Housing Economics 9*, 197-213.

PAINTER, G. YANG, L. and YU, Z. (2004): "Homeownership determinants for Chinese Americans: assimilation, ethnic concentration and nativity", *Real Estate Economics 32 (3)*, 509-539.

PAINTER, G. and GABRIEL, S. (2008): "Mobility, residential location and the American dream: the intra-metropolitan geography of minority homeownership", *Real Estate Economics 36 (3)*, 499-531.

PERACCHI, F. (2002): "The European Community Household Panel: A Review", Empirical Economics 27, 63-90.

SMITS and MULDER (2008): "Family dynamics and first-time homeownership", *Housing Studies 23*, 917-933.

ZORN, P. (1988): "An analysis of household mobility and tenure choice: an empirical study of Korea", *Journal of Urban Economics 24*, 113-128.

Table 1: Descriptive statistics

	Model	*Individual data*		*Person-period data*	
Variables		*Mean*	*Std. dev.*	*Mean*	*Std. dev.*
Transition	Change tenure status	0.404	0.491	0.354	0.478
Tenure status	Tenure status	0.877	0.329	0.902	0.297
Stayer	Selection eq.sequential	0.968	0.174		
Mover	Selection eq.recent-movers	0.524	0.499		
Tenure status$_{t-1}$	Markov model	0.872	0.334		
Gender	All	0.780	0.414	0.787	0.410
Permanent income	All	2.655	1.111	2.638	1.110
Age	All	56.011	16.529	59.941	14.968
Rooms	Change tenure status	5.183	1.163	5.259	1.168
Relative price	All	5335.192	21996.10	6213.349	24907.55
Education	All				
University level		0.160	0.367	0.134	0.341
Secondary education		0.123	0.328	0.102	0.303
Primary education		0.717	0.451	0.763	0.425
Household size	Tenure status	3.195	1.532	3.056	1.326
Change in household size	Change tenure status	0.033	0.516	0.061	0.520
Employment status					
Change in employment status				0.027	0.161
Working week > 15 hrs	Tenure status	0.478	0.492		
Working week < 15 hrs	Tenure status	0.009	0.093		
Unemployed	Tenure status	0.041	0.198		
Long-term unemployed	Tenure status	0.003	0.050		
Inactive	Tenure status	0.461	0.498		
Marital status					
Change in marital status	Change tenure status			0.002	0.049
Married	Tenure status	0.712	0.472		
Separated	Tenure status	0.022	0.146		
Divorced	Tenure status	0.014	0.116		
Widowed	Tenure status	0.160	0.366		
Single	Tenure status	0.083	0.277		
Duration	Change in tenure status	14.269	6.313	16.984	4.256

Table 2: Significance of the coefficients of the different models

	Classic	*Markov*	*Recent-movers*	*Sequential*	*Duration*	*Choice-based*
Income	+	+	+	+	+	+
Previous tenure	-	+	-	-	-	-
Sex	0	-	0	0	0	0
Age	+	0	+	+	+	+
Age^2	+	+	0	+	+	+
Studies	*0*	0	0	0	0	0
Employment situation	+	0	+	0	0	0
Members	*0*	+	0	0	+	+
Marital status	+	0	+	+	+	+
Relative price	+	0	0	+	+	+
Number of Rooms	-	-	-	-	+	+
Duration	-	-	-	-	+	+

Table 3a: Predictive ability of models with tenure status as dependent variable

	Classical model	*Markov model*	*Recent-movers model*	*Sequential model*
Renting	0.05	0.76	0.04	0.15
Ownership	0.99	0.98	0.99	0.98
Total	0.88	0.96	0.87	0.80

Table 3b: Predictive ability of models with transition from renting to ownership as dependent variable

	Duration model	*Choice-based model*
No transition	0.62	0.68
Transition	0.56	0.65
Total	0.62	0.68

Table 4a: Marginal effects in models where the dependent variable is tenure status

	Classical model		*Markov model*		*Recent-movers model*		*Sequential model*	
Variables	*Marginal effect*	*%*	*Marginal effect*	*%*	*Marginal effect*	*%*	*Marginal effect*	*%*
Probability (tenure status=1)	0.902		0.860		0.841		0.815	
Income	1.48*	1.64	1.23 *	1.43	3.93*	4.67	2.30*	2.82
Gender	0.53	0.59	-	-	-3.48	-4.14	-0.89	-1.09
Age	1.06*	1.17	-3.28	-3.82	1.40*	1.66	1.93*	2.37
Age^2	-0.01*	-0.01	0.02*	0.02	-0.01	-0.01	-0.02*	-0.02
Education (ref.: university level)								
Secondary education	0.62	0.68	0.65	0.75	1.92	2.29	0.41	0.50
Primary education	3.53*	3.92	2.14*	2.48	5.81*	6.90	3.38*	4.14
Employment status (ref.: week > 15 hrs)								
Working week < 15h	-9.93	-11.01	10.48	12.19	-19.00*	-22.59	8.33	10.21
Unemployed	-5.64	-6.25	-0.94	-1.09	-8.18*	-9.72	-4.09	-5.01
Long-term unemployed	-9.17	-10.17	1.91	2.23	-4.33*	-5.16	8.58	10.51
Inactive	-2.74	-3.04	-1.04	-1.22	3.46*	4.11	-3.14	-3.85
Marital status (ref.: married)								
Separated	-10.90*	-12.08	-32.69*	-38.02	-15.06*	-17.91	-20.01*	-24.54
Divorced	-7.39*	-8.19	-8.74*	-10.16	-15.13*	-17.99	-22.65*	-27.77
Widowed	-4.03	-4.47	2.70*	3.14	-6.18	-7.35	-6.46	-7.92
Single	-4.73	-5.24	-14.80*	-17.21	-12.77*	-15.18	-5.23	-6.41
Household size	0.46*	0.51	-4.08*	-4.75	-0.36	-0.43	0.09	0.10
Relative price	0.11*	0.12	-0.03*	-0.03	0.01	0.01	-1.04*	-1.28
Tenure status $_{t-1}$			78.42*	91.20				
N	4026		4026		4026		4026	

* Significant (p-value < 0.05)

Table 4b: Marginal effects in models where the dependent variable is the transition of tenure status (from renting to ownership)

	Duration model		*Choice-based model*	
Variables	*Marginal effects (%)*	*%*	*Marginal effects (%)*	*%*
Probability (transition=1)	*0.0019*		*0.0019*	
Income	0.05*	24.36	0.03*	17.10
Gender	0.03	17.44	0.00	5.26
Age	0.01*	5.64	0.03*	17.00
Age^2	0.00*	0.00	0.00*	0.01
Education (ref.: university level)				
Secondary education	0.05	24.63	0.00	6.53
Primary education	0.05*	27.70	0.01*	36.84
Employment status	-0.01	-3.47	0.01	7.20
Household size	0.08*	43.08	0.04*	21.05
Marital status	0.25*	126.57	0.15*	73.68
Bedrooms	-0.01*	-6.67	-0.04*	-19.15
Duration	-0.01*	-3.08	-0.05*	-27.36
Relative price	0.00*	0.75	0.00*	0.93
N	44472		4712	

* Significant (p-value < 0.05)