Draft

Tenure Choice with Sample Selection: A Note on the Differences among Alternative Samples

by

Gary Painter*
Assistant Professor
School of Policy, Planning, and Development
University of Southern California
Los Angeles, California 90089-0626

February 2000

* The author thanks Stuart Gabriel, Dowell Myers, and seminar participants at the Lusk Center for Real Estate at USC, as well as participants at the 2000 AREUEA annual conference for valuable comments. The author also acknowledges the Mortgage Bankers Association of America and the Zumberge Research and Innovation Fund of the University of Southern California for research funding, and Paul Hughes for his excellent research assistance.

Tenure Choice with Sample Selection:
A Note on the Differences among Alternative Samples

by

Gary Painter*
Assistant Professor
School of Policy, Planning, and Development
University of Southern California
Los Angeles, California 90089-0626
gpainter@usc.edu
phone: (213) 740-8754
fax: (213) 740-0001

Abstract

Most tenure choice models using cross sectional data have used either a sample of recent movers or a sample comprising all households. There are problems with estimating both types of models in cross sectional data. A sample of recent movers over-samples renters, and a sample of all households will yield estimates based on household decisions made in the past. This research designs a method to correct for sample selection in a sample of recent movers. There are large differences in the importance of age and immigrant status and duration as a predictor of homeownership. At the same time, income effects are similar across models.

JEL classification: R21

Keywords: Tenure choice, Housing Demand

I. Introduction

Recent years have witnessed substantial academic and practitioner research regarding the determinants of homeownership. One of the reasons for this research is that by understanding the determinants of homeownership, one may be able to formulate policy that encourages higher levels of ownership. This is undertaken with the belief that homeownership attainment generates neighborhood benefits as pertains to property upkeep, public safety, school quality, and the like (see, for example, Green and White, 1997).

In the literature, there have been two primary methods that have used to estimate tenure choice models. Most models employ the use of cross-sectional data. The first method used by Ihlanfeldt (1981) and others uses a sample of recent movers. The rationale behind using this group is that the decisions of recent movers are more likely to reflect equilibrium conditions in the housing market. At the same time, using a sample of recent movers may be biased if there are no controls for previous tenure and because renters are over-represented in the sample. For example, Boehm, Herzog, and Schlottmann (1991) use longitudnal data to determine previous tenure.

The other approach to modeling tenure choice has been to estimate the model using a sample of all households (Edin and Englund 1991, Wachter and Megbolugbe 1992, Gyourko and Linneman 1996, and Coulson 1999). This approach models cumulative homeownership attainment. The rationale for this approach is that homeownership is a long term decision based as much upon anticipated future needs as on present needs (Wachter and Megbolugbe, 1992: 339; Edin and Englund, 1991). However, as Painter, Gabriel, and Myers (1999) note, among households who are age 45

or older, cumulative attainment of homeownership may largely reflect the lagged effects of past choices. Therefore, this approach, if used in cross-sectional data, is biased as current data is not likely to reflect past choices among homeowners.

An alternative approach in cross-sectional data is to use a sample of recent movers, but to account for the probability that someone is likely to be a mover. This approach has the advantage of not having the lagged effect of tenure choice models among all households and therefore is more likely to capture tenure choice in equilibrium conditions. At the same time, most cross-sectional data does not account for previous tenure status. In order to properly account for the likelihood that someone is a mover, this analysis employs a Heckman-style (1979) correction in which a first step probit is estimated capturing the decision to move. Because both dependent variables are binary, a bivariate probit model with sample selection is proposed. Estimation of tenure choice models among movers and among all households is compared to the selection-corrected model in order to determine the importance of selection correction.

Results suggest that controlling for the likelihood that someone is a mover in the selection model has important impacts on the coefficients of the tenure choice model. While education and income variables have stable coefficients across models and samples, the other estimated effects differ by model specification. In particular, the age of the householder is much less important in the selection model than in the other approaches. Also, status as an immigrant is not as large of a detriment to homeownership once the likelihood of moving is included in the model.

II. Data

Data used in this analysis are drawn from the public use microdata sample (PUMS) file of the 1990 decennial census. This is a 5% sample of all households living in Los Angeles County, which constitutes the Los Angeles-Long Beach primary metropolitan statistical area.[1] The data provide detailed information about both the housing unit, and the individuals who reside in that unit.

The full sample consists of 96,548 households. This sample includes all households which either own or rent their primary residence, excluding households which reside in group quarters. This sample includes household heads that are aged 18-64, because the elderly may have significantly different tenure choice behavior. In addition, the analysis is restricted to four racial/ethnic categories for which there is sufficient sample size for stratified analysis: white, non-Hispanic; black, non-Hispanic; Hispanic, non-Asian; and Asian. This is done to test for the sensitivity of estimates across groups. Of the full sample, 52,656 had moved within the past five years. It is these households which form the movers sample, although the full sample is used in both the estimation of the tenure attainment equations and in the estimation of the decision to move.

The complete list of variables selected for analysis and their definitions are given in Appendix 1. The independent variables include demographic factors (race-ethnicity, age, marital status, number of people in the household, number of workers in the household, migrant origin and history), economic (salary income, dividend and other income, education level of the householder), and other factors which affect housing tenure choice.

Like most other studies, wealth effects cannot be measured directly with the data at hand. As such, the educational attainment of the household head is used as a proxy to

[1] There are a total of 295,489 households in 1990.

indicate the future earnings potential as well as the wealth of the family. Presumably, movers with higher levels of human capital are more capable of meeting downpayment requirements.

As is evident in the housing literature, proxies for the relative costs of owning to renting and household income are fundamental to economic models of housing tenure choice (e.g., Ihlanfeldt, 1981). Unlike most studies, which utilize a national sample of observations, the present study uses data from on a single metropolitan area. Therefore, intra-metropolitan variations in house prices or rents are not distinguished; instead, households are assumed to face the same rent and price frontier within the metropolitan area. This specification is consistent with recent additions to the tenure choice literature (see, for example, Wachter and Megbolugbe [1992], Gyourko and Linneman [1996], and Coulson [1999]), which used metropolitan level variation house prices and rents to identify those effects. Evidence from Painter, Gabriel, and Myers (1999) suggests that this is a reasonable assumption in a single metropolitan area. Therefore, controls for household income proxy the effects of nominal housing affordability on household tenure choice. In that regard, both permanent and transitory measures of household income are in the tenure choice equation. Using the method of Goodman and Kawai (1982), permanent income is the predicted value of a regression of household income on a set of demographic and human capital characteristics.[2] Transitory income is calculated as the residual of observed household income and predicted income.

The analysis further adjusts for immigrant status and history (interacted with ethnicity and by years in the U.S. since immigration) as well as migrant origin (entered as a series of categorical variables indicating whether the household moved from within Los

Angeles County, moved from elsewhere in the U.S., or moved from outside the U.S.). Controlling for immigration timing and ethnicity, newcomers to a region may have lower homeownership probabilities than do longer-term residents. Newcomers by definition are mobile and are more often drawn from the ranks of renters. Migrants may also undertake extensive search prior to investment in housing. Also, relative to local homeowners who may have benefited from substantial house price appreciation, migrants may be characterized by more binding homeownership affordability constraints.

Table 1 presents the means of the independent variables used in the study for both the whole sample and the movers only sample.[3] The reference household is chosen to be white, married, aged 25-34, with a high school diploma, and a non-immigrant who has moved from within Los Angeles County. While most of the characteristics are similar across movers and non-movers, there are three primary differences. First, movers are more likely to be younger and not married. Second, movers have lower levels of permanent income and slightly more transitory income. Finally, immigrants are more likely to be movers. This is truer for those with the smallest length of stay in the U.S. To the extent that these differences are important, then estimating models with movers, and with both non-movers and movers could lead to quite different results.

III. Empirical Model

For both the models used most commonly in the literature, authors typically use a logit or a probit specification to estimate the binary tenure choice model. The only difference is that one group (e.g. Wachter and Megbolugbe [1992], Gyourko and Linneman [1996], and Coulson [1999]) uses all households in the sample, and the other

[2] Results of these household income regressions are available upon request.
[3] Tables of means by race/ethnicity status are available from the author upon request.

group (Ihlanfeldt, 1981) uses a sample of movers. Boehm, Herzog, and Schlottmann (1991) present a multinomial logit model of mobility and tenure choice, but they used longitudinal data, and estimation of their model is not possible in many cross-sectional data sets.

In this note, a model of tenure choice with sample selection is introduced. When estimating a model of movers, a household's choice of tenure is not observed if they do not move. Therefore, standard estimation of tenure choice among movers is biased. Sample selection bias is accounted for by employing a variant of Heckman's (1979) two-step selection model. The model of tenure choice among movers which corrects for selection bias is adapted from Van de Ven and Van Pragg (1981), in which both the selection equation and the tenure choice equation have binary dependent variables.[4] (Boyes, Hoffman, and Low [1989] presents a similar econometric model applied to the problem of simultaneously estimating default and the application process for credit cards when default is only observed for the sample of applicants for which credit is approved.)

As with the standard formulation, assume that there exists a latent variable OWN* that measures the propensity to own among mover households in the sample. The observable tenure choice indicator is regressed on a vector of demographic, economic, and other factors affecting the housing tenure decision.

$$OWN^{*}_{i} = X_i \beta + \varepsilon_{1i}$$

such that one observes only the binary outcome,

[4] The two-step selection model is often estimated by obtaining the inverse Mill's Ratio from a first stage probit, and then entering it into the second stage equation. Edin and Englund (1991) present this type of approach in an appendix. As noted by van de Van and van Praag (1981), if the dependent variable in the second stage equation is binary, the error term does not have a normally distributed error term; and therefore a two-stage approach for this problem would yield only approximate results. The difference between the two approaches in this case is confined to reducing the coefficients on the age coefficients, but the remainder of the coefficients of the model are unchanged.

$OWN_i = 1$, if $OWN^*_i > 0$ and
$OWN_i = 0$, if $OWN^*_i \# 0$.

However, one only observes OWN_i for observation *i* if $MOVE_i = 1$, where $MOVE^*_i$ is taken from the underlying relationship,

$MOVE^*_i = Z_i$ (+ $,_{2i}$, where
$MOVE_i = 1$, if $MOVE^*_i > 0$ and
$MOVE_i = 0$, if $MOVE^*_i \# 0$.

Finally, the assumption is made that $,_{1i}$, and $,_{2i}$ are jointly normally distributed with correlation coefficient D. This allows maximum likelihood estimation of the log likelihood function

$$L = \sum_{i \in S}^{yi=1} \ln[\Phi_2(\mathrm{Xi}\ b, \mathrm{Z_i}\, g, r)] + \sum_{i \in S}^{yi=0} \ln[\Phi_2(-\mathrm{Xi}\ b, \mathrm{Z_i}\, g, r)] + \sum_{i \notin S} \ln[1 - \Phi_1(\mathrm{Z_i}\, g)]$$

where *S* is the set of observations for which OWN_i is observed (recent movers), M_1 is the standard cumulative normal and M_2 is the cumulative bivariate normal distribution function. Unlike the standard Heckman selection model, the bivariate probit with sample selection is weakly identified without the use of identifying assumptions in the selection equation (Greene, 1997). Likelihood ratio tests confirm that they are not necessary.

IV. Results

Regression coefficients and their standard errors from two probit models of housing tenure choice in alternate sample and the model with sample selection are displayed in Table 2 for the unified sample; estimation results for each of the race-ethnicity stratifications are contained in Tables 3-6. (Appendix 2 shows the estimates of the selection equation in the unified sample.) In Table 2, the coefficients have the expected signs. For example, both permanent and transitory income increases the

probability of homeownership, and not having a high school diploma lowers the probability of homeownership. Across models, there are many differences in the estimated effects, but there is also some stability in the coefficients. In particular, the impact of income and education is similar across models.

On the other hand, there are some important differences as evidenced in Table 2. First, the importance of the age of the household differs markedly across the models. In particular, the tenure choice model with the full sample (Column 1) attributes a large positive effect for older households. This is reduced in a model of movers (Column 2), and is further reduced when the probability that a household moves is included in the model (Column 3). For example, the coefficient on a household aged 45-54 is 0.608 in the tenure status model, is 0.258 in a model of movers, and is only 0.078 in the model with sample selection. This implies that once controls for economic factors are included, being older does not predict as much of a higher homeownership rate as previously thought. In addition, the negative impact of being unmarried is less for movers than for the full sample.

The other major differences are in the variables capturing race/ethnicity, immigrant and domestic migrant status. In either model of movers (Columns 2 & 3), the penalty for not having resided in Los Angeles in the previous 5 years is smaller than in the tenure status model (Column 1). For example, the coefficient on moved from with the U.S., but outside California, is close to -0.7 for both models with movers, but is about -0.9 in the model with all households.

Also evidenced in Table 2 is the fact that there is a larger penalty among movers for being black or Latino, and a smaller benefit for being Asian. With regard to

immigrant status, being an immigrant is much less of a detriment to homeownership in the model with sample selection (Column 3) than in the other models. Further, the duration of residence, which is a positive determinant of homeownership, also is less important in the model with sample selection.

Tables 3 through 6 present results stratified by race/ethnicity because Wachter and Megbolugbe (1992) and Painter, Gabriel, and Myers (1999) have shown that coefficients can differ substantially by race/ethnicity. For whites (Table 3), the pattern of results is much the same as in the unified sample. The economic and human capital factors differ little by model specification, and the age and domestic migrant status differ considerably. The lone exception is for immigrants. While immigrant status is a smaller detriment to homeownership (similar to unified sample), there is little difference in the estimated coefficients on duration of residence. This may be due to the fact that there are relatively few white immigrants. As is evidenced in Table 4, the pattern for whites is similar to the pattern for blacks. The sole difference concerns the impact of immigrants, as it is positive for blacks.

The basic pattern of results is replicated in the Latino sample (Table 5), but not in the Asian sample (Table 6). In the selection-corrected model (Column 3) and to a lesser extent in the model of movers without selection correction (Column 2), the younger households have higher homeownership than do the older households. Also, in the selection-corrected model, being an immigrant is a positive predictor of homeownership. This is in sharp contrast to Coulson (1999) and the results from the model with the full sample (Column 1) which both suggested that being an immigrant had a substantial negative effect on homeownership for Asians.

Finally, the statistical significance of the correlation coefficient suggests that controlling for the likelihood of moving is critical to determining the effects of the socioeconomic characteristics on tenure choice. The only sample in which the correlation coefficient is not significant is the sample of black households, meaning that there is no difference between the model of movers with or without the correction for sample selection. For the remainder of the samples, as documented above, correction for the likelihood of moving has an important impact on the interpretation of the determinants of tenure choice.

V. Conclusion

This note has demonstrated the importance of correcting for sample selection in the estimation of tenure choice models when using cross sectional data. Other models overestimate the importance of age and immigrant status on tenure choice. The bias is greatest in the models that utilize the full sample of households. Others have recognized the potential for these problems and have attempted to lower potential bias by estimating separate models by age category (e.g., Gyourko and Linneman 1996) or including previous tenure in the equation, but this will not eliminate the bias on the other coefficients of the model. At the same time, the importance of economic and human capital factors appears to be invariant between models. This result provides confidence that importance of permanent income and transitory income is robust across samples and techniques.

The results here with respect to the importance of the age of the householder are very similar to longitudnal studies such as Boehm, Herzog, and Schlottmann (1991), as well as to the age cohort approach advocated by Myers, Megbolugbe, and Lee (1998).

The Boehm, Herzog, and Schlottmann (1991) study found that age was not important for tenure choice, but was for the decision to move. Also similar to the results here, Myers, Megbolugbe, and Lee (1998) found that after a household reached age 25, there was little difference in the homeownership propensities of these age groups. These two other approaches are sensible when the data allow for their estimation, but in cross-sectional data, this note suggests that it is prudent to control for the likelihood of moving to obtain accurate estimates of the determinants of tenure choice.

References

Boehm, T., H. Herzog, and A. Scholottmann (1991): "Intra-Urban Mobility, Migration, and Tenure Choice," *The Review of Economics and Statistics*, 73(1), 59-68.

Boyes, W., D. Hoffman, and S. A. Low(1989): "An Econometric Analysis of the Bank Credit Scoring Problem," *Journal of Econometrics*, 40, 3-14.

Coulson, E. (1999): "Why are Hispanic- and Asian-American Homeownership Rates So Low?: Immigration and Other Factors," *Journal of Urban Economics*, 45, 209-227.

Edin, P-A. and P. Englund (1991): "Moving Costs and Housing Demand: Are Recent movers really in Equilibrium?" *Journal of Public Economics*, 44, 299-320.

Goodman, A. C. (1988): "An econometric model of housing price, permanent income, tenure choice and housing demand," *Journal of Urban Economics*, 23, 327-354.

Goodman, A. C. and M. Kawai (1982): "Permanent Income, Hedonic Prices and Demand for Housing: New Evidence," *Journal of Urban Economics*, 12, 214-37.

Green, R. K. and M. J. White (1997): "Measuring the Benefits of Homeowning: Effects on Children," *Journal of Urban Economics*; 41 (3), 441-61.

Gyourko, J. and P. Linneman (1996): "An analysis of the changing influences on traditional household ownership patterns," *Journal of Urban Economics*, 39, 318-341.

Heckman, J. J. (1979): "Sample selection bias as a specification error," *Econometrica*, 47, 153-161.

Ihlanfeldt, K. (1981): "An empirical investigation of alternative approaches to estimation of the equilibrium demand for housing," *Journal of Urban Economics*, 9, 97-105.

Myers, D., I. Megbolugbe, and S. W. Lee (1998): "Cohort estimation of homeownership attainment among native-born and immigrant populations," *Journal of Housing Research*, 9(2), 237-269.

Painter, G., S. Gabriel, and D. Myers (1999): "Race, Immigrant Status, and Housing Tenure Choice," Lusk Center for Real Estate Working Paper, 99-09.

Van de Ven, W. P. M. M. and B. M. S. Van Pragg (1981): "The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection," *Journal of Econometrics*, 17, 229-252.

Wachter, S. and I. Megbolugbe (1992): "Racial and Ethnic Disparities in Homeownership," *Housing Policy Debate*, 3, 333-370.

Table 1
Variable Summary Statistics

	N = 96548		N = 52656	
	Mean	Std Dev.	Mean	Std Dev.
OWNERSHIP RATE	0.517	0.500	0.373	0.484
AGE 18-24	0.051	0.221	0.083	0.275
AGE 25-34	0.269	0.443	0.390	0.488
AGE 35-44	0.286	0.452	0.295	0.456
AGE 45-54	0.214	0.410	0.148	0.356
AGE 55-64	0.180	0.384	0.084	0.277
NOT MARRIED MALE HEAD OF HOUSEHOLD	0.204	0.403	0.243	0.429
NOT MARRIED FEMALE HEAD OF HOUSEHOLD	0.258	0.437	0.265	0.441
LESS THAN A HIGH SCHOOL DEGREE	0.172	0.378	0.167	0.373
HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE	0.417	0.493	0.406	0.491
COLLEGE DEGREE OR BETTER	0.411	0.492	0.427	0.495
NUMBER OF PEOPLE IN HOUSEHOLD	2.930	1.735	2.876	1.744
NUMBER OF WORKERS IN HOUSEHOLD	1.699	0.985	1.656	0.932
PERMANENT INCOME (1000s)	52.335	25.934	48.619	25.159
TRANSITORY INCOME (1000s)	0.000	38.165	0.127	37.025
ETHNICITY- WHITE	0.591	0.492	0.582	0.493
ETHNICITY- BLACK	0.119	0.324	0.111	0.314
ETHNICITY- LATINO	0.172	0.377	0.170	0.375
ETHNICITY- ASIAN	0.118	0.322	0.137	0.344
MOVED FROM WITHIN CALIFORNIA	0.048	0.213	0.067	0.250
MOVED FROM WITH U.S	0.069	0.253	0.126	0.332
MOVED FROM A FOREIGN COUNTRY	0.042	0.201	0.078	0.268
IMMIGRANT	0.283	0.451	0.319	0.466
LATINO IMMIGRANT	0.100	0.300	0.106	0.307
ASIAN IMMIGRANT	0.096	0.295	0.119	0.324
CAME TO U.S WITHIN 5 YEARS	0.048	0.213	0.081	0.273
CAME TO U.S 5-10 YEARS AGO	0.062	0.241	0.084	0.278
CAME TO U.S 10-15 YEARS AGO	0.056	0.229	0.063	0.243
CAME TO U.S 15-20 YEARS AGO	0.039	0.193	0.036	0.186
CAME TO U.S 20-30 YEARS AGO	0.049	0.215	0.037	0.190
CAME TO U.S MORE THAN 30 YEARS AGO	0.030	0.172	0.018	0.132

Note: Variable definitions are provided in Appendix 1.

Table 2 - Tenure Choice Among Different Models

	(1) Full Sample of Households		(2) Movers Sample - No Selection Correction		(3) Movers Sample - Selection Correction	
VARIABLE	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
INTERCPT	-0.759	0.030	-0.931	0.040	-1.026	0.040
AGE 18-24	-0.314	0.030	-0.415	0.034	-0.386	0.034
OMITTED: AGE 25-34						
AGE 35-44	0.301	0.013	0.149	0.017	0.057	0.019
AGE 45-54	0.608	0.016	0.258	0.022	0.078	0.029
AGE 55-64	0.873	0.016	0.344	0.027	0.088	0.037
NOT MARRIED, MALE HEAD OF HOUSEHOLD	-0.560	0.016	-0.442	0.022	-0.406	0.022
NOT MARRIED, FEMALE HEAD OF HOUSEHOLD	-0.406	0.019	-0.344	0.027	-0.319	0.026
OMITTED: MARRIED						
NO HIGH SCHOOL DIPLOMA	-0.205	0.016	-0.191	0.023	-0.179	0.022
OMITTED: HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE						
COLLEGE DEGREE OR BETTER	0.020	0.017	-0.032	0.023	-0.017	0.023
NUMBER OF PEOPLE IN HOUSEHOLD	0.048	0.004	0.031	0.005	0.024	0.005
NUMBER OF WORKERS IN HOUSEHOLD	-0.131	0.011	-0.192	0.015	-0.184	0.015
PERMANENT INCOME (1000s)	0.019	0.001	0.022	0.001	0.021	0.001
TRANSITORY INCOME (1000s)	0.011	0.000	0.011	0.000	0.011	0.000
ETHNICITY- BLACK	-0.241	0.017	-0.311	0.026	-0.323	0.025
ETHNICITY- LATINO	0.030	0.019	-0.006	0.028	-0.029	0.028
ETHNICITY- ASIAN	0.249	0.033	0.243	0.046	0.199	0.045
OMITTED: WHITE						
MOVED FROM WITHIN CALIFORNIA	-0.501	0.023	-0.301	0.023	-0.293	0.022
MOVED FROM WITHIN U.S	-0.910	0.022	-0.719	0.022	-0.703	0.021
MOVED FROM A FOREIGN COUNTRY	-0.648	0.041	-0.553	0.042	-0.540	0.041
OMITTED: MOVED FROM WITHIN LA CO.						
IMMIGRANT	-0.476	0.044	-0.186	0.049	-0.034	0.051
IMMIGRANT*ASIAN	-0.004	0.040	0.035	0.053	0.079	0.052
IMMIGRANT*LATINO	-0.277	0.029	-0.178	0.040	-0.193	0.039
OMITTED: BORN IN THE US						
CAME TO U.S 5-10 YEARS AGO	0.089	0.042	0.162	0.045	0.082	0.045
CAME TO U.S 10-15 YEARS AGO	0.389	0.043	0.421	0.047	0.297	0.049
CAME TO U.S 15-20 YEARS AGO	0.574	0.045	0.497	0.052	0.349	0.054
CAME TO U.S 20-30 YEARS AGO	0.703	0.045	0.528	0.054	0.361	0.056
CAME TO U.S MORE THAN 30 YEARS AGO	0.706	0.050	0.395	0.065	0.214	0.067
OMITTED: CAME TO U.S. IN THE PAST 5 YEARS						
Correlation Coefficient (Ð)					0.300	0.032
Log Likelihood Function	-47813		-26158		-81617	
Number of Households	96548		52656		52656	
Mean of Dependent Variable	0.517		0.373		0.373	

Table 3 - Tenure Choice Among Different Models
Sample of White Households

	(1) Full Sample of Households		(2) Movers Sample - No Selection Correction		(3) Movers Sample - Selection Correction	
VARIABLE	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
INTERCPT	-0.660	0.040	-0.787	0.053	-0.825	0.054
AGE 18-24	-0.372	0.039	-0.455	0.044	-0.440	0.045
OMITTED: AGE 25-34						
AGE 35-44	0.319	0.018	0.157	0.022	0.111	0.027
AGE 45-54	0.667	0.021	0.304	0.029	0.218	0.042
AGE 55-64	0.902	0.021	0.363	0.034	0.240	0.055
NOT MARRIED, MALE HEAD OF HOUSEHOLD	-0.606	0.021	-0.517	0.028	-0.504	0.028
NOT MARRIED, FEMALE HEAD OF HOUSEHOLD	-0.447	0.025	-0.391	0.035	-0.383	0.035
OMITTED: MARRIED						
NO HIGH SCHOOL DIPLOMA	-0.212	0.024	-0.177	0.034	-0.171	0.034
OMITTED: HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE						
COLLEGE DEGREE OR BETTER	0.025	0.022	-0.033	0.029	-0.026	0.029
NUMBER OF PEOPLE IN HOUSEHOLD	0.088	0.006	0.053	0.008	0.048	0.008
NUMBER OF WORKERS IN HOUSEHOLD	-0.166	0.015	-0.248	0.021	-0.246	0.021
PERMANENT INCOME (1000s)	0.016	0.001	0.021	0.001	0.020	0.001
TRANSITORY INCOME (1000s)	0.010	0.000	0.010	0.000	0.010	0.000
MOVED FROM WITHIN CALIFORNIA	-0.524	0.026	-0.325	0.027	-0.324	0.026
MOVED FROM WITHIN U.S	-0.930	0.025	-0.743	0.025	-0.739	0.025
MOVED FROM A FOREIGN COUNTRY	-0.838	0.069	-0.693	0.070	-0.689	0.069
OMITTED: MOVED FROM WITHIN LA CO.						
IMMIGRANT	-0.384	0.074	-0.138	0.079	-0.065	0.084
OMITTED: BORN IN THE US						
CAME TO U.S 5-10 YEARS AGO	0.162	0.084	0.270	0.090	0.226	0.091
CAME TO U.S 10-15 YEARS AGO	0.384	0.082	0.404	0.090	0.345	0.092
CAME TO U.S 15-20 YEARS AGO	0.370	0.093	0.331	0.106	0.269	0.108
CAME TO U.S 20-30 YEARS AGO	0.516	0.082	0.321	0.094	0.244	0.098
CAME TO U.S MORE THAN 30 YEARS AGO	0.551	0.082	0.311	0.098	0.227	0.102
OMITTED: CAME TO U.S. IN THE PAST 5 YEARS						
Correlation Coefficient (D)					0.131	0.047
Log Likelihood Function	-28248		-15959		-48482	
Number of Households	57105		30672		30672	
Mean of Dependent Variable	0.572		0.418		0.418	

Table 4 - Tenure Choice Among Different Models
Sample of Black Households

	(1) Full Sample of Households		(2) Movers Sample - No Selection Correction		(3) Movers Sample - Selection Correction	
VARIABLE	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
INTERCPT	-1.318	0.069	-1.415	0.100	-1.504	0.119
AGE 18-24	-0.095	0.090	-0.399	0.125	-0.383	0.126
OMITTED: AGE 25-34						
AGE 35-44	0.462	0.042	0.266	0.056	0.194	0.090
AGE 45-54	0.892	0.046	0.444	0.069	0.289	0.168
AGE 55-64	1.265	0.047	0.619	0.081	0.418	0.218
NOT MARRIED, MALE HEAD OF HOUSEHOLD	-0.544	0.050	-0.367	0.074	-0.327	0.086
NOT MARRIED, FEMALE HEAD OF HOUSEHOLD	-0.421	0.054	-0.344	0.083	-0.320	0.085
OMITTED: MARRIED						
NO HIGH SCHOOL DIPLOMA	-0.133	0.041	-0.073	0.069	-0.065	0.068
OMITTED: HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE						
COLLEGE DEGREE OR BETTER	0.131	0.050	0.074	0.073	0.072	0.072
NUMBER OF PEOPLE IN HOUSEHOLD	0.002	0.010	-0.036	0.015	-0.039	0.015
NUMBER OF WORKERS IN HOUSEHOLD	-0.029	0.035	-0.090	0.054	-0.089	0.054
PERMANENT INCOME (1000s)	0.020	0.002	0.023	0.003	0.022	0.003
TRANSITORY INCOME (1000s)	0.014	0.001	0.013	0.001	0.013	0.001
MOVED FROM WITHIN CALIFORNIA	-0.538	0.118	-0.211	0.110	-0.210	0.108
MOVED FROM WITHIN U.S	-0.967	0.089	-0.677	0.087	-0.667	0.088
MOVED FROM A FOREIGN COUNTRY	-0.858	0.246	-0.689	0.245	-0.676	0.241
OMITTED: MOVED FROM WITHIN LA CO.						
IMMIGRANT	0.101	0.279	0.498	0.265	0.590	0.280
OMITTED: BORN IN THE US						
CAME TO U.S 5-10 YEARS AGO	-0.602	0.318	-0.687	0.313	-0.739	0.312
CAME TO U.S 10-15 YEARS AGO	-0.168	0.311	-0.321	0.311	-0.373	0.312
CAME TO U.S 15-20 YEARS AGO	-0.041	0.315	-0.146	0.324	-0.223	0.333
CAME TO U.S 20-30 YEARS AGO	-0.192	0.300	-0.332	0.303	-0.401	0.308
CAME TO U.S MORE THAN 30 YEARS AGO	-0.050	0.338	-0.450	0.428	-0.515	0.429
OMITTED: CAME TO U.S. IN THE PAST 5 YEARS						
Correlation Coefficient (Ð)					0.231	0.231
Log Likelihood Function	-5501		-2274		-9183	
Number of Households	11486		5834		5834	
Mean of Dependent Variable	0.368		0.196		0.196	

Table 5 - Tenure Choice Among Different Models
Sample of Latino Households

	(1) Full Sample of Households		(2) Movers Sample - No Selection Correction		(3) Movers Sample - Selection Correction	
VARIABLE	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
INTERCPT	-0.828	0.063	-1.017	0.083	-1.138	0.087
AGE 18-24	-0.405	0.071	-0.501	0.080	-0.466	0.079
OMITTED: AGE 25-34						
AGE 35-44	0.267	0.033	0.117	0.043	0.046	0.043
AGE 45-54	0.547	0.039	0.219	0.058	0.081	0.057
AGE 55-64	0.856	0.042	0.373	0.075	0.164	0.069
NOT MARRIED, MALE HEAD OF HOUSEHOLD	-0.510	0.042	-0.408	0.057	-0.379	0.056
NOT MARRIED, FEMALE HEAD OF HOUSEHOLD	-0.409	0.048	-0.470	0.071	-0.446	0.070
OMITTED: MARRIED						
NO HIGH SCHOOL DIPLOMA	-0.167	0.032	-0.203	0.046	-0.201	0.045
OMITTED: HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE						
COLLEGE DEGREE OR BETTER	-0.046	0.049	-0.008	0.067	0.019	0.067
NUMBER OF PEOPLE IN HOUSEHOLD	0.044	0.007	0.040	0.010	0.036	0.009
NUMBER OF WORKERS IN HOUSEHOLD	-0.162	0.028	-0.189	0.039	-0.180	0.038
PERMANENT INCOME (1000s)	0.023	0.002	0.024	0.003	0.023	0.003
TRANSITORY INCOME (1000s)	0.015	0.001	0.015	0.001	0.015	0.001
MOVED FROM WITHIN CALIFORNIA	-0.378	0.072	-0.146	0.071	-0.142	0.069
MOVED FROM WITHIN U.S	-0.864	0.100	-0.610	0.095	-0.598	0.093
MOVED FROM A FOREIGN COUNTRY	-0.441	0.094	-0.358	0.093	-0.352	0.091
OMITTED: MOVED FROM WITHIN LA CO.						
IMMIGRANT	-0.904	0.091	-0.603	0.109	-0.490	0.109
OMITTED: BORN IN THE US						
CAME TO U.S 5-10 YEARS AGO	0.082	0.087	0.183	0.096	0.126	0.095
CAME TO U.S 10-15 YEARS AGO	0.419	0.085	0.563	0.096	0.471	0.095
CAME TO U.S 15-20 YEARS AGO	0.787	0.086	0.845	0.100	0.724	0.100
CAME TO U.S 20-30 YEARS AGO	0.986	0.088	0.948	0.106	0.820	0.106
CAME TO U.S MORE THAN 30 YEARS AGO	1.017	0.101	0.789	0.135	0.658	0.135
OMITTED: CAME TO U.S. IN THE PAST 5 YEARS						
Correlation Coefficient (D)					0.271	0.030
Log Likelihood Function	-7832		-3854		-13638	
Number of Households	16609		8937		8937	
Mean of Dependent Variable	0.401		0.264		0.264	

Table 6 - Tenure Choice Among Different Models
Sample of Asian Households

	(1) Full Sample of Households		(2) Movers Sample - No Selection Correction		(3) Movers Sample - Selection Correction	
VARIABLE	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
INTERCPT	-0.726	0.081	-1.071	0.102	-1.376	0.094
AGE 18-24	-0.056	0.090	-0.113	0.099	-0.093	0.091
OMITTED: AGE 25-34						
AGE 35-44	0.090	0.039	0.042	0.044	-0.092	0.043
AGE 45-54	0.122	0.047	-0.039	0.058	-0.299	0.060
AGE 55-64	0.244	0.051	-0.046	0.070	-0.384	0.076
NOT MARRIED, MALE HEAD OF HOUSEHOLD	-0.286	0.051	-0.172	0.061	-0.137	0.056
NOT MARRIED, FEMALE HEAD OF HOUSEHOLD	-0.089	0.060	-0.046	0.076	-0.011	0.070
OMITTED: MARRIED						
NO HIGH SCHOOL DIPLOMA	-0.187	0.048	-0.161	0.061	-0.121	0.057
OMITTED: HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE						
COLLEGE DEGREE OR BETTER	0.000	0.050	-0.059	0.062	-0.038	0.057
NUMBER OF PEOPLE IN HOUSEHOLD	0.042	0.010	0.035	0.012	0.018	0.011
NUMBER OF WORKERS IN HOUSEHOLD	-0.151	0.031	-0.160	0.038	-0.150	0.036
PERMANENT INCOME (1000s)	0.025	0.002	0.027	0.003	0.025	0.003
TRANSITORY INCOME (1000s)	0.013	0.001	0.012	0.001	0.010	0.001
MOVED FROM WITHIN CALIFORNIA	-0.457	0.065	-0.348	0.064	-0.295	0.058
MOVED FROM WITHIN U.S	-0.700	0.061	-0.622	0.060	-0.555	0.055
MOVED FROM A FOREIGN COUNTRY	-0.533	0.067	-0.488	0.070	-0.454	0.065
OMITTED: MOVED FROM WITHIN LA CO.						
IMMIGRANT	-0.387	0.081	-0.010	0.096	0.438	0.099
OMITTED: BORN IN THE US						
CAME TO U.S 5-10 YEARS AGO	0.111	0.064	0.156	0.070	-0.019	0.069
CAME TO U.S 10-15 YEARS AGO	0.439	0.069	0.413	0.078	0.088	0.082
CAME TO U.S 15-20 YEARS AGO	0.571	0.076	0.355	0.090	-0.064	0.095
CAME TO U.S 20-30 YEARS AGO	0.591	0.084	0.352	0.106	-0.179	0.112
CAME TO U.S MORE THAN 30 YEARS AGO	0.727	0.113	0.169	0.170	-0.440	0.165
OMITTED: CAME TO U.S. IN THE PAST 5 YEARS						
Correlation Coefficient (Ρ)					0.672	0.069
Log Likelihood Function	-5721		-3842		-9828	
Number of Households	11348		7213		7213	
Mean of Dependent Variable	0.557		0.459		0.459	

Appendix 1

Variable Definitions

Throughout, the unit of observation is the head of household. Those aged less than 18 years, or greater than 65 years, have been excluded. Also excluded are certain racial categories, mainly American Indians. This is done to ensure that the four racial/ ethnic groups (black, Asian, Latino and white) constitute the entire sample used in the analysis. In all the regressions, only those people who lived in Los Angeles county at the time of the census, and in a different house, which may or may not have been in Los Angeles county, five years before the census was taken.

AGE 18-24	People aged 18 through 24 inclusive.
OMITTED CATEGORY: AGE 25-34	People aged 25 through 34 inclusive.
AGE 35-44	People aged 35 through 44 inclusive.
AGE 45-54	People aged 45 through 54 inclusive.
AGE 55-64	People aged 55 through 64 inclusive.
NOT MARRIED MALE HEAD OF HOUSEHOLD	Head of household is male, and is not married (i.e.; he is divorced, separated, never married or widowed).
NOT MARRIED FEMALE HEAD OF HOUSEHOLD	Head of household is female, and is not married (i.e.; she is divorced, separated, never married or widowed).
OMITTED CATEGORY: MARRIED	Head of household is married, and is not separated.
NO HIGH SCHOOL DIPLOMA	High school not completed, or not yet.
OMITTED CATEGORY: HS DIP/NO COL DEGREE	High school completed, but not four years of post-high school education.
COLLEGE DEGREE OR BETTER	Minimum of four years of post-high school education is completed.
NUMBER OF PEOPLE IN HOUSEHOLD	This number includes people of all ages, including those aged less than 18 years and 65 or older.
NUMBER OF WORKERS IN HOUSEHOLD	A worker is defined as somebody who worked in the year before the census was conducted.
WAGE&SALARY	Wage and salary income aggregated across all members of the household.

WAGE&SALSQ	Wage and salary income squared aggregated across all members of the household.
DIV INCOME	Dividend, interest and rental income aggregated across all members of the household.
OTHINCOME	All other types of income aggregated across all members of the household. That is non-farm self-employment income, farm income, social security income, public assistance income, retirement income (only in 1990- this category does not exist in 1980) and other income.
ETHNICITY- BLACK	Black, non-Hispanic.
ETHNICITY- ASIAN	Asian, which may be Hispanic.
ETHNICITY- LATINO	Hispanic, non-Asian.
OMITTED CATEGORY: WHITE	White, non-Hispanic.
MOVED FROM WITHIN CALIFORNIA	People who lived in California five years ago, but not in Los Angeles county. i.e.; this would include someone who moved from San Francisco to Pasadena, but not someone who moved from Santa Monica to Pasadena.
MOVED FROM WITHIN U.S	People who lived in the United States five years ago, but not in California. i.e.; this would include someone who moved from Oklahoma to Los Angeles, but not from San Francisco to Los Angeles.
MOVED FROM A FOREIGN COUNTRY	People who lived outside the United States five years ago.
OMITTED CATEGORY: MOVED FROM LA CO.	People who lived in Los Angeles county five years ago.
IMMIGRANT	Someone who is not a citizen of the U.S, or is a citizen, but only by naturalization. A non-immigrant is thus someone who was born in the US, Puerto Rico, Guam and outlying areas or who was born abroad of American parents.

IMMIGRANT*ASIAN	This takes the value of 1 if a person is both Asian and an immigrant, as defined above, and zero otherwise.
IMMIGRANT*LATINO	This takes the value of 1 if a person is both Latino and an immigrant, as defined above, and zero otherwise.
OMITTED CATEGORY: CAME TO U.S IN THE PAST 5 YEARS	An immigrant, as defined above, who arrived within 5 years of the taking of the census.
CAME TO U.S 5-10 YEARS AGO	An immigrant, as defined above, who arrived between 5 and 10 years before the taking of the census.
CAME TO U.S 10-15 YEARS AGO	An immigrant, as defined above, who arrived between 10 and 15 years of the taking of the census.
CAME TO U.S 15-20 YEARS AGO	An immigrant, as defined above, who arrived within 15 and 20 years of the taking of the census.
CAME TO U.S 20-30 YEARS AGO	An immigrant, as defined above, who arrived within 20 and 30 years of the taking of the census.
CAME TO U.S MORE THAN 30 YEARS AGO	An immigrant, as defined above, who arrived more than 30 years before the taking of the census.

Appendix 2 - Results of the Selection Equation
Dependent Variable = Moved in the Past 5 Years

VARIABLE	Coef.	Std. Error
INTERCPT	0.789	0.028
AGE 18-24	0.283	0.027
OMITTED: AGE 25-34		
AGE 35-44	-0.607	0.013
AGE 45-54	-1.056	0.015
AGE 55-64	-1.404	0.016
NOT MARRIED, MALE HEAD OF HOUSEHOLD	0.176	0.016
NOT MARRIED, FEMALE HEAD OF HOUSEHOLD	0.092	0.019
OMITTED: MARRIED		
NO HIGH SCHOOL DIPLOMA	0.033	0.015
OMITTED: HIGH SCHOOL DIP. BUT NO COLLEGE DEGREE		
COLLEGE DEGREE OR BETTER	0.108	0.016
NUMBER OF PEOPLE IN HOUSEHOLD	-0.040	0.004
NUMBER OF WORKERS IN HOUSEHOLD	0.031	0.011
PERMANENT INCOME (1000s)	-0.002	0.001
TRANSITORY INCOME (1000s)	0.000	0.000
ETHNICITY- BLACK	-0.121	0.017
ETHNICITY- LATINO	-0.146	0.019
ETHNICITY- ASIAN	-0.246	0.030
OMITTED: WHITE		
IMMIGRANT	1.296	0.038
IMMIGRANT*ASIAN	0.287	0.037
IMMIGRANT*LATINO	-0.103	0.029
OMITTED: BORN IN THE US		
CAME TO U.S 5-10 YEARS AGO	-0.811	0.038
CAME TO U.S 10-15 YEARS AGO	-1.079	0.038
CAME TO U.S 15-20 YEARS AGO	-1.207	0.040
CAME TO U.S 20-30 YEARS AGO	-1.296	0.040
CAME TO U.S MORE THAN 30 YEARS AGO	-1.377	0.044
OMITTED: CAME TO U.S. IN THE PAST 5 YEARS		
Correlation Coefficient (D)	0.300	0.032
Number of Households	96548	
Mean of Dependent Variable	0.545	