Housing Tenure Choice and the Dual Income Household

Steven Carter *
Department of Economics
University of California, Irvine

November 9, 2006

Abstract

This paper presents an econometric model to estimate the effect of a household's secondary income on the tenure choice of housing. The model gives immediate results as to the effect of interest, but also determines the endogeneity of secondary income in the same procedure. Results show that there is a substantial effect of secondary income on tenure choice, and that it is endogenous. Households that have a positive secondary income can be 32% more likely to own a dwelling than those who have no secondary income.

1 Introduction

This paper answers two question concerning a household's tenure choice of housing:

1. "What impact does a secondary income in the household have on the decision to own or rent a dwelling?"

2. "Is the secondary income endogenous to the tenure choice of the household?"

With a drastic increase in housing prices across the country over the past several years and no comparative increase in wages, it would seem that the only way households would be able to afford a house[1] is by sending the non-working spouse into the work force. If the choice of sending the non-working spouse into the work force is endogenous, then it would

Address for correspondence: Department of Economics, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100. E-mail: sjcarter@uci.edu.

[1]This paper will use a loose definition of "house". A house will be considered any dwelling that can be owned. This can include single family detached homes, condominiums or apartments. This definition is required because the PSID only asks if the respondent owns the dwelling.

be expected that the choice is based on certain covariates concerning the current state of the household (i.e. the current household income, education level of the non-working spouse). However, to qualify financially for a mortgage, the total household income is more important the the number of incomes in the household. For those households on the margin (meaning the margin of whether or not to send the non-working spouse into the workforce *and* not owning a home), the decision would need to take into account whether or not the added income would not only be enough to qualify for a mortgage, but cover the costs of lost home production. Since the amount of the secondary income is more important in terms of financial qualification for a mortgage, this paper focuses mainly on testing the impact of the dollar amount of the secondary income on tenure choice and it's endogeneity, rather than just its presence in the household.

The questions presented above are answered through a Bayesian simultaneous equation model with limited dependent variables. This modelling is first proposed by Li (1998). Li presents the model in terms of a Tobit/Probit framework. This paper adapts her model and algorithm to fit the above questions. One of the advantages of this modelling is that the parameter that answers the endogeneity question is readily available after one estimation routine, where as the classical method would require a two step procedure.

The paper proceeds as follows: Section 2 reviews the current discussion of the housing tenure choice literature, specifically how it related to the dual income nature of the household. Section 3 describes the regression model in the tenure choice context and estimation algorithm. Section 4 discusses the data and estimation results and Section 5 concludes.

2 Housing Tenure Choice

Housing tenure choice is a topic of thorough study, but also a rich topic of study. As there can be many factors influencing the decision to own a home, there is a breadth of literature on the different factors[2]. Since the scope of the present paper concentrates on secondary

[2]There is the tax arbitrage theory, which examines the trade off between renting and owning a dwelling based on the tax bracket of the household (Englund and Persson (1982), Follian and Ling (1988), Narwold

income (additional labor supply), a look into the labor economics literature provides some insight into this topic. This area focuses on the conditional state of housing consumption and it's affect on labor supply. Given a particular mortgage constraint, Fortin (1995) shows that a significant percentage of households increase the female contribution of labor supply. Aldershof et al (1997) examine the household maximization of utility over housing consumption, female leisure and all other goods, where housing consumption and labor supply are jointly determined. Both of these studies treat the measure of housing consumption (mortgage payment or other similar measure) as exogenous and out of the hands of the household. Bottazzi (2004) tests the endogeneity of the mortgage commitment and finds that it is endogenous. However, these studies treat the causation direction as tenure commitment leads to participation of a second labor supplier.

In areas where there is high housing inflation, the labor supply decision may affect the tenure choice (instead of the tenure choice affecting labor supply). Households may choose to send a second laborer into participation based on the high user cost of housing for the purpose of saving enough to make a down payment or afford potentially high monthly payments. Based on this possibility, modelling the decision alone to supply labor to the market may not capture the effect of that decision as well as looking at the amount that the secondary earner makes. If the secondary earner has a low potential wage, then a household may not choose to purchase. However, if the potential wage is high enough, then the household may commit to a mortgage based on the amount of that additional income. Also, since this is a decision of the household, the variable may be endogenous and should be modelled appropriately. This study contributes by providing an empirical test of the endogeneity of secondary income and it's effects on housing tenure choice.

and Sonstelie (1994)). High bracket households will buy housing and become landlords for tax shelter purposes, and rent the dwelling to low bracket households for a price lower than the cost of owning a home. There is the idea of considering home ownership as part of the investment portfolio as well as a consumption good. (Brueckner (1997), Flavin and Yamashita (2001), Fratantoni (1998)) This literature shows sub-optimal investment portfolios exist if demand for housing exceeds the mean-variance efficient quantitiy of housing. Empirical tests of the covariance theory of income and rents are available in Davidoff (2006) and Ortalo-Magné and Rady (2002), which show that high covariances between income and housing prices decrease the value of owner-occupied housing.

3 The Model

To measure empirically the impact of secondary income on the tenure choice of housing, a choice model can be specified to incorporate the secondary income as well as other appropriate variables used in the previous literature. However, if secondary income is considered endogenous, then the secondary income must be modelled as it's own regression equation.

3.1 The Two-Equation Specification

It can argued that the effect of a secondary income on the tenure choice of a household is endogenous. In the extreme case, a household can choose to be a two income household for the sole purpose of owning the dwelling. The secondary income would allow for better financial circumstances to qualify for financing, or to place a bigger down payment. If this were the case, then the contribution of the secondary income to the tenure choice would be correlated with the error term in the Probit regression, leading to biased and inconsistent results. To correct for this, the secondary income is modelled as a simultaneously estimated Tobit regression with the Probit regression.

The Tobit regression for secondary income is modelled as follows:

$$y_2^* = X_2\beta_2 + \epsilon_2 \tag{1}$$

y_2^* is the latent variable as described in Chib (1992). For the context of this paper, the latent variable will be the net income (total income minus cost of labor participation). Under this description, observations of no income have found that the cost of working is higher than the salary received. $y_2^* = y_2$ if $y_2^* > 0$, where $+y_2$ is the observed data and $y_2^* = 0$ otherwise. ϵ_2 is a normally distributed error term with a zero mean and variance σ_{22}. Together the equations make the system

$$y_1^* = y_2\gamma_2 + X_1\beta_1 + \epsilon_1 \tag{2}$$

$$y_2^* = X_2\beta_2 + \epsilon_2 \tag{3}$$

The covariate matrix X_1 are variables that affect housing tenure choice. The covariate matrix

X_2 includes those variables in X_1 as well as an instrument to make the system identifiable. Together, ϵ_1 and ϵ_2 have a bivariate distribution with mean $\mathbf{0}$ and covariance matrix Σ defined as

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \right).$$

The unity restriction is preserved in the covariance matrix, Σ, where the first element $\sigma_{11} = 1$. When choosing the priors for the Gibbs sampler, this unity restriction slightly complicates matters, in that a standard Inverse Wishart prior cannot be used. The Drèze and Richard decomposition can be used to place priors on the individual elements of Σ. Priors for the entire system are discussed in the following section.

There are a few items worth noting about the system. First, as mentioned in Li (1998), the system is conditional on the latent variables (already augmented) y_1^* and y_2^*. Also, given the latent data, the system has a Seemingly Unrelated Regression (SUR) format which will simplify the data augmentation within the Gibbs sampler.

3.2 Simultaneous Equation Priors

Prior distributions for the regression parameters above are proper conjugate priors. Since the model above has the same format as a SUR model, all the slope coefficients are given the prior $(\gamma_2, \beta_1, \beta_2) \sim MVN(\beta_0, \Psi_0)$.

As mentioned earlier, an Inverse Wishart prior cannot be used on Σ due to the unity restriction. However, Drèze and Richard (1983) show that, given a fixed matrix Q (in this case, Q is the updating matrix of errors from the SUR model) the covariance matrix can be decomposed into the following product of independent distributions:

$$P(\Sigma|Q) = P(\sigma_{11}|Q_{11}, \nu - 2)P(\sigma_{11}^{-1}\sigma_{12}|\sigma_{22.1})P(\sigma_{22.1}|Q_{22.1}, \nu) \tag{4}$$

where $\sigma_{22.1} = \sigma_{22} - \sigma_{11}^{-1}\sigma_{12}^2$. In the case of Σ for this paper, by the unity restriction, $P(\sigma_{11})$=1. The remaining elements of the decomposition have the following distributions:

$$P(\sigma_{11}^{-1}\sigma_{12}|\sigma_{22.1}) \sim N(Q_{11}^{-1}Q_{12}, \sigma_{22.1}Q_{11}^{-1}) \tag{5}$$

and

$$P(\sigma_{22.1}|Q_{22.1}) \sim IG(Q_{22.1}, \nu) \tag{6}$$

3.3 Two-Equation Algorithm

This section describes the estimation algorithm for the Simultaneous Equation model. The algorithm consists of three updating steps for the estimation parameters and two data augmentation steps for the limited dependent variables. The algorithm is very similar to Li (1998), but varies in the updating for the covariance matrix (Li's unity restriction is placed on the second equation, where this paper places the unity restriction on the first). The algorithm is as follows:

1. Conditional on y_1^*, y_2^* and Σ

$$\begin{pmatrix} \gamma_2 \\ \beta_1 \\ \beta_2 \end{pmatrix} \sim N(\bar{\beta}, \bar{\Psi}) \tag{7}$$

   where $\bar{\Psi} = Z'(\Sigma \otimes I_n)Z + \Psi_0$ and $\bar{\beta} = \Psi(Z'(\Sigma \otimes I_n)Y^* + \Psi_0\beta_0)$. Note that $\bar{\Psi}$ is the covariance matrix, and not the precision.

2. Conditional on y_1^*, y_2^* and parameters $(\gamma_2, \beta_1, \beta_2)$,

$$\sigma_{22.1} \sim IG(Q_{22.1}, \nu) \tag{8}$$

   where $\nu = \nu_0 + n$ and

$$Q_{22.1} = Q_{22} - Q_{11}^{-1}Q_{12}^2 \tag{9}$$

   where $Q_{22} = (y_2^* - X_2\beta_2)'(y_2^* - X_2\beta_2)$, $Q_{11} = (y_1^* - y_2\gamma_2 - X_1\beta_1)'(y_1^* - y_2\gamma_2 - X_1\beta_1)$ and $Q_{12} = (y_1^* - y_2\gamma_2 - X_1\beta_1)'(y_2^* - X_2\beta_2)$.

3. Conditional on y_1^*, y_2^*, parameters $(\gamma_2, \beta_1, \beta_2)$ and $\sigma_{22.1}$,

$$\sigma_{12} \sim N(Q_{11}^{-1}Q_{12}, \sigma_{22.1}Q_{11}^{-1}) \tag{10}$$

   where Q_{11} and Q_{12} are defined above. Note that once $\sigma_{22.1}$ and σ_{12} are drawn, σ_{22} can be backed out from the definition of $\sigma_{22.1}$.

4. Conditional on y_2^*, parameters $(\gamma_2, \beta_1, \beta_2)$ and Σ

$$y_1^* \sim TN(\mu_{1|2}, \sigma 1|2) \tag{11}$$

where TN represents the truncated normal distribution. $\mu_{1|2} = X_1\beta_1 + \frac{\sigma_{12}}{\sigma_{22}}(y_2^* - X_2\beta_2)$ and $\sigma_{1|2} = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}$. y_1^* is drawn from the normal truncated above zero if $y_1 = 0$ and drawn from the normal truncated below zero is $y_1 = 1$.

5. Conditional on y_1^*, parameters $(\gamma_2, \beta_1, \beta_2)$ and Σ

$$y_2^* \sim TN(\mu_{2|1}, \sigma_{2|1}). \tag{12}$$

Since y_2 is only censored below zero, y_2^* is only drawn from the normal truncated above zero when $y_2 = 0$. $\mu_{2|1} = X_2\beta_2 + \frac{\sigma_{12}}{\sigma_{11}}(y_1^* - y_2\gamma_2 - X_1\beta_1)$ and $\sigma_{2|1} = \sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}$. If y_2¿0, then there is no need for data augmentation. In the context of this paper, the drawing of y_2^* below zero simulates the household's negative net income from the labor participation of the secondary worker, meaning that there is some cost to the household that, if the secondary worker chooses to participate in the labor force, then the costs of participation exceed the income earned.

The three Gibbs sampler steps and the two data augmentation steps complete the estimation algorithm. This algorithm has been tested with artificial data for accuracy and recovers parameter estimates quite well. As mentioned above, the simulation of σ_{12} is identical to a classical approach to a two stage estimation procedure, as the updating formulae for σ_{12} are the same as regressing the errors from the first equation on the errors of the second equation. Inference about endogeneity is then made on the σ_{12} parameter.

4 Data and Results

4.1 Data

The data come from the 2001 wave of the Panel Study of Income Dynamics (PSID), a rich data set that surveys households about household income, household characteristics and

other socio-demographic variables. A benefit from these data is that both the tenure choice of the household and the spouses income are available. The PSID, however, does not indicate the type of dwelling in which the household resides, as the type of dwelling was shown to be a significant modelling choice (Brownstone and Englund, 1991). The sample only includes married households, since the interest lies in the secondary income contribution to the tenure choice. The list of variables, with their sample means and standard deviations, are shown in Table 1.

The secondary income variable is defined as the minimum of the head of household and wife income variables, as there might be some households where the wife is the primary bread-winner and the husband provides the household production. As total and taxable income are highly co-linear, this paper employs disposable income as the measure of household income. Disposable income is the difference of total household income less taxable income (a linear function of the two). Another important variable is the user cost of housing (UCH). This variable, computed at the individual level, measures the housing expenditures per month of the household. If the household is a renter, the UCH is the monthly rent. For owner-occupiers, the UCH is the monthly mortgage payment plus the monthly property tax value of the property. Unfortunately, this variable does not have much meaning for those who would be thinking of changing tenure, as the variable captures the individual UCH as opposed to an aggregate UCH for a geographic area[3]. The metropolitan indicator is a dummy variable indicating if the household lives in a metropolitan area with a population of at least 250,000 people. The average age of couple variable is the basic average of the head of household and wife's ages. The highest level of education variable is defined as the maximum of the head of household's and wife's education levels. There is also a variable indicating whether or not that educational level was obtained by the head of household.

In considering endogeneity, proper instruments are needed to rectify the existing correlation between the endogenous variable and the error term. This paper uses the age of the youngest

[3]At the current time, the author is waiting for approval to access sensitive data from the PSID. Once granted, the data will allow the author to make census level UCH calculations for a better measure.

Table 1: Means and Standard Deviations of Variables, N=3136

Variable	Mean	St. Dev.
Tenure choice (1=own)	0.71	0.45
Secondary income amount	16963	17962
Secondary income amount (excluding zero earners)	23219	17215
Total household disposable income	8191	22461
Metro indicator (1=Area Population greater than 250,000)	0.67	0.47
Average age of couple	41.7	18.6
Average age of youngest child	4.35	5.27
Notable percentages		
Single income households	26.9	
Zero disposable income	44.9	
No children in the household	38.4	

child as the instrument to identify the secondary income variable in the first equation. The age will affect the choice of the secondary earner's labor force participation in that older children may not need to be supervised and the stay-at-home parent can work. However, age of the youngest child does should not affect the tenure choice of the household.

4.2 Results

The parameters of interest are the effect of secondary income (γ_2) and the endogeneity of that variable (σ_{12}). Since the conditional distributions of these parameters are known, inference can be made from the simulated draws of these distributions. The results presented below are estimated with non-informative priors, though with the large sample size, even Gibbs sampler runs with misleading priors (i.e. prior values have opposite signs from the results from the non-informative case) converge to the same results.

Table 2 presents the results from the regression parameters. Since the question at hand deals with housing tenure choice, Table 2 shows results of the variables affecting housing tenure choice. It also gives results of the instrument used to identify the system, age of youngest child. Though not used directly in the tenure choice regression, the age has a positive effect on the secondary income variable from the Tobit regression.

Secondary income increases the probability of household ownership, holding disposable in-

Table 2: Conditional Means of Parameter Distributions, 10000 draws

Variable	Mean	St. Dev.	Median	Mode
Constant	-1.12	0.37	-1.12	-1.24
Secondary Income	0.17	0.02	0.17	0.16
Disposable Income	0.03	0.01	0.03	0.03
User Cost of Housing	0.00	0.00	0.00	0.00
Average age of couple	0.01	0.00	0.01	0.01
Number of Kids	0.05	0.02	0.05	0.06
North east	-0.16	0.41	-0.17	0.33
North Central	-0.12	0.41	-0.13	-0.19
South	-0.19	0.41	-0.20	-0.36
West	-0.22	0.41	-0.22	-0.25
Metro Dummy	-0.13	0.05	-0.13	-0.15
Age youngest	0.14	0.02	0.14	0.14

come constant. Just from the cross section alone, statements about the dynamics of the household decision to supply more labor in order to own the housing cannot be made, but the positive effect shows that there is some relationship between the amount of secondary income and the ownership decision. Though the values in Table 2 are not marginal effects, some probability evaluations are made later. Table 3 provides 99th percentile intervals for the conditional distributions for each variable. Lower and upper bounds of the same sign show that the bulk of the distribution is centered away from a zero effect.

Table 3: 99th percentile intervals of parameter distributions, $\mu \pm 2.576\sigma$

Variable	Lower bound	Upper bound
Constant	-2.08	-0.17
Secondary Income	0.13	0.21
Disposable Income	0.01	0.04
User Cost of Housing	0.00	0.00
Average age of couple	0.01	0.01
Number of Kids	0.00	0.09
North east	-1.22	0.89
North Central	-1.18	0.94
South	-1.23	0.86
West	-1.26	0.83
Metro Dummy	-0.25	0.00
Age youngest	0.10	0.18

Inference of the parameter σ_{12} gives information regarding the endogeneity of secondary

income. In the classical setting, a significant difference from zero would indicate that the variable in question is endogenous. In this setting, the centering of the distribution away from zero would indicate that secondary income is endogenous to housing tenure choice. Table 4 gives the results of the distribution of σ_{22}, which indicates that secondary income is endogenous to the housing tenure choice and that failure to treat it as such leads to biased estimates of it's effects.

Table 4: Conditional Mean of σ_{12} from 10000 draws, and 99th percentile interval $\mu \pm 2.576\sigma$

	Mean	St. Dev.	Median	Mode	Lower bound	Upper bound
σ_{12}	-2.78	0.30	-2.79	-2.83	-3.56	-2.01

Since the results from the probit regression are not marginal effects, the coefficient on secondary income does not reveal much about the relative probability of choosing one alternative over the other. The relative probability by varying the secondary income comes from evaluation of the normal cumulative distribution function for each observation and then averaging over the number of iterations. So, the probability of choosing ownership, given the regression parameters, is

$$P(y_{1i} = 1|\beta_1, \beta_2, \gamma_2, \Sigma) = 1 - \Phi(-(y_{2i}\gamma_2 + X_{1i}\beta_1 + \epsilon_2\sigma_{12})).$$

For this paper, the difference was evaluated for households with zero secondary income and \$16000 secondary income. Since the probability is evaluated for every iteration of the Gibbs sampler, the probability of choosing ownership based on this difference of secondary income is the average of each evaluation of the probability above. For households with \$16000 secondary income, the probability of ownership is .96 (standard deviation of .06), where those with no secondary income is .64 (standard deviation of .20) yielding a difference of .32. This result includes the condition that households have the same level of disposable income, meaning that the marginal contribution of the primary earner's income for the dual income household is less than the household with only one income. An explanation for this result might be that households with one income have higher, unobservable home production costs

(Becker, (1964)) than households with two incomes and save more by only supplying one laborer to the market.

5 Conclusion

This paper shows that consideration of the secondary income of a household is a substantial gain in estimating of the housing tenure choice factors of a household. Also, the is evidence that this variable is endogenous and failure to recognize this leads to biased estimates. From the estimation model presented here, the probability of a household with a secondary income of $16000 is 32% higher than a household with no secondary income, all else equal. Though the estimation method is sound, the control for the user cost of housing is weak and should be reconsidered pending approval to access better data.

References

Aldershof, T., R. Alessie and A. Kapteyn (1997) "Female labor supply and the demand for housing" Progress Report 46, CentER, Tillburg University

Becker, G.S. (1964) "Human Capital" (Columbia University Press for the National Burea, of Economic Research, New York)

Bottazzi, R. (2004) "Labour market participation and mortgage borrowing constraints" *Institute for Fiscal Studies* Working Paper.

Brownstone, D. and P. Englund (1991) "The Demand for Housing in Sweden: Equilibrium Choice of Tenure and Type of Dwelling" *Journal of Urban Economics*, 29, 267-281.

Brueckner, J. (1997) "Consumption and investment motives and portfolio choices of homeowners" *Journal of Real Estate Finance and Economics*, 15, 159180.

Chib, S. (1992) "Bayes inference in the Tobit censored regression model", *Journal of Econometrics*, 51, 79-99.

Davidoff, T. (2006) "Labor income, housing prices and homeownership" *Journal of Urban Economics*, 59, 209-235.

Englund, P. and M. Persson (1982) "Housing prices and tenure choice with asymmetric taxes and progressivity", *Journal of Public Economics*, 19, 271-290.

Flavin, M. and T. Yamashita (2001) "Owner-occupied housing and the composition of the household portfolio over the life-cycle", *American Economic Review*

Follian, R. and D. C. Ling, (1988) "Another looks at thenure choice, inflation and taxes", *AREUEA Journal*, 16, 207-229.

Fortin, N. (1995) "Allocation inflexibilities, Female labor supply and housing assets accumulation: Are women working to pay the mortgage?", *Journal of Labor Economics*,

13, 524-557.

Fratantoni, M. (1998) "Homeownership and investment in risky assets", *Journal of Urban Economics*, 44, 2742.

Li, K. (1998) "Bayesian inference in simultaneous equation models with limited dependent variables", *Journal of Econometrics*, 85, 387-400.

Narwold A. and J. Sonstelie (1994), 'State income taxes and homeownership: a test of the tax arbitrage theory", *Journal of Urban Economics*, 36, 249-277.

Ortalo-Magné, F. and S. Rady, "Tenure choice and the riskiness of non-housing consumption", *Journal of Housing Economics*, 11, 266279.