WORKING PAPER

ITLS-WP-14-07

Housing prices and price endogeneity in tenure and dwelling type choice models

By
Chinh Ho and David Hensher

April 2014

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in
Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-14-07

TITLE: **Housing prices and price endogeneity in tenure and dwelling type choice models**

ABSTRACT: The application of a strategic transport – land use model usually requires dwelling price data for both the chosen and non-chosen alternatives. Previous studies have used average dwelling prices either externally sourced from Census and commercial data or internally generated by hedonic models. The use of average prices may be reasonable if the spatial resolution is fine enough to ensure the absence of price endogeneity– a well known issue but hardly investigated in the literature on housing choices. Using data collected from a revealed preference survey conducted in Sydney in 2013, this paper presents a hedonic price model to obtain price data for each dwelling as a function of dwelling and location characteristics. The data are then used as input into a tenure and dwelling type choice model and a test undertaken to investigate the possible presence of price endogeneity, which we show to not exist. Price elasticities of demand for different dwelling types are derived and the application of this model within a broader and very general framework of the new version of the 'Transportation and Environment Strategy Impact Simulator' (New-TRESIS) is discussed.

KEY WORDS: ***Dwelling prices, hedonic price, choice models, endogenity, integrated transport and land use models***

AUTHORS: **Ho and Hensher**

Acknowledgements: This paper is part of a large project at ITLS developing a new transportation and location strategic modelling system (new TRESIS or TRANSCAN). ITLS team members are Michiel Bliemer, David Hensher, Chinh Ho, Mark Raadsen, Richard Ellison and Jun Zhang.

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (C13)

The Australian Key Centre in Transport and Logistics Management

The University of Sydney NSW 2006 Australia

Telephone: +612 9114 1824

E-mail: business.itlsinfo@sydney.edu.au

Internet: http://sydney.edu.au/business/itls

DATE: April 2014

1. Introduction

Dwelling price is one of the most important factors in explaining housing choice, be it the decision relating to residential location, tenure type or dwelling type. However, the dwelling price is unavailable in most datasets, especially for non-chosen alternatives. For modelling purposes, research on housing choices has to estimate this critical information for use over a population of dwellings. Two methods have been used in the literature: one uses average sales price and renting cost, and the other uses a hedonic modelling technique to generate price data for non-chosen alternatives. Both methods of imputing price data have potential issues which lead to a biased parameter estimate for dwelling price. On the one hand, the use of an average dwelling price may not represent different market values faced by different households in the same area. This issue is pronounced if a rough spatial classification is used in aggregating the average dwelling price, which is the case for most previous studies. On the other hand, using a hedonic function to generate dwelling prices may result in price endogeneity in which a certain feature of the dwelling is factored into its price but is not included in the model.

The literature has recognised that decisions on residential location, tenure type and dwelling type are interrelated. Given a limited housing budget, households may choose to purchase a small unit in the inner area or to rent a large detached dwelling in a fringe area, depending on their preference. The challenge is to set up a practical modelling system which takes the interdependencies between the choices of residential location, tenure and dwelling type into account.

This paper contributes to the literature by developing a survey instrument and hedonic model to generate dwelling prices and weekly renting costs for modelling the choice of tenure and dwelling type in which price endogeneity is tested. The paper provides evidence on the price elasticities of demand for different dwelling types within a tenure category in Sydney. The next section reviews the literature on dwelling choices. This is followed by a description of the survey and modelling approach. Estimation results of dwelling prices and household choice of tenure and dwelling type choice are then presented. The application of the model within a general framework of the TRESIS system is discussed and a way of capturing interdependencies between dwelling choices and between land use and transport is proposed. The paper ends with a summary of the main findings.

2. Literature review

The literature on housing choices covers three key household decisions: residential location, tenure and dwelling type. Although these decisions are interrelated, most studies separate them into sequential choices for model estimation and practical implementation. A utility maximisation framework can then be employed to link these choice models together in a sequential order with interrelated choices being captured through constraints and logsums (or expected maximum utility), usually in a nested logit structure. This section reviews housing choice studies with the focus on dwelling and tenure type decisions, as well as methodologies for modelling these discrete choices, which provide a context for this paper.

One of the earliest discrete choice models capturing the interdependencies between tenure and dwelling type choices was developed by Boehm (1982) who used a hierarchical model to encompass three dimensions of household choices including tenure (own vs. rent), location (high vs. low income neighbourhoods) and dwelling type (large vs. small dwellings). The model was structured and estimated in a hierarchical manner with the highest level being the tenure choice, followed by location and dwelling type choices. The study found that household size, lifecycle, race, prior dwelling value, and dwelling price were important to housing choices. In addition, the effect of income was significant and different across dwelling and tenure types, but not so across residential locations. One main concern about the hierarchical modelling approach is that interdependencies between choices are captured in only one direction through conditional probability with no feedback from the lower level to the preceding level in the hierarchy through logsum or maximum expected utility as in a genuine nested logit model. A main reason for using hierarchical models was that they were easier to estimate than multinominal logit (MNL) and nested logit models (Boehm, 1982).

Using data from the Australian census in 1986 and 1996, Yates and Mackay (2006) examined different choice structures in modelling all three choices of residential location (inner vs. outer area), tenure type (rent vs. own) and dwelling type (house vs. flat). They compared estimation results of an MNL model with those of NL models and concluded that the structure of the error terms, and consequently the modelling technique, was important to providing meaningful results. They also found that dwelling price, household income and household structure significantly influence housing choices.

Numerous studies have focused on two dimensions of housing choice, with the most common being tenure choice (Cho, 1997; Rapaport, 1997; Skaburskis, 1999; Tu and Goldfinch, 1996). Most of these studies relied on the workhorse of MNL models when it comes to estimating the joint choice of tenure and dwelling type/location, without careful consideration of the error structure as to whether assumptions on the distribution of error terms were satisfied. Examples were Cho (1997) and Skaburski (1999) who both investigated the joint choice of dwelling and tenure type. Although the rationale for the use of a genuine NL model was explicitly recognised by Cho and Skaburski, they used MNL models due to the limitation of employed software and estimation technique at the time. Both justified the use of MNL model based on an argument coined by Tu and Goldfinch (1996) who suggested that the choice of modelling approach has little effect on the final outcome if households have full market information when making their choice. However, as Yates and Mackay (2006) have pointed out, this argument is oblivious of the fact that the violation of the *iid* assumption of the error terms underlying MNL models comes from the limited capability of the analyst, as opposed to whether the consumer has full information.

Rapaport (1997) also used an MNL model but expanded the joint choice of location and tenure type to include the quantity of housing, measured as the ratio of the market value to the price of a standard unit of housing. In her model, each household in the local market faced a choice amongst 10 combinations of 5 communities and 2 tenure types. She used two hedonic models, one for renters and one for owners, to form a dwelling price index for each of the dwellings in the sample. The estimated dwelling price index showed a strong and significant influence on the joint choice of location and tenure type, consistent with results from other studies (Cho, 1997; Tu and Goldfinch, 1996; Yates and Mackay, 2006). In addition, the study found that government expenditure, school quality, household permanent income and lifecycle were important to location and tenure choices.

A common element in all but one of these studies is the use of average dwelling prices or a relative price index for non-chosen alternatives, without correcting (if it exists) price endogeneity. If some structural or spatial characteristics of the dwelling are partially factored into its price but are not included in the model, then estimation without regard to the correlation between price and unobserved factors is biased. The exception is the study by Rapaport (1997) who used a two-step procedure to obtain consistent estimates of dwelling price for the quantity of housing consumed. However, this study, as do many others, uses a rough classification of neighbourhoods (e.g., inner vs. outer) and thus, average prices may not represent the heterogeneous prices that different households in the same location face with. In both cases of price endogeneity and price heterogeneity, the estimation of price elasticity of demand for housing is inconsistent.

This paper addresses the issue associated with supplementing prices for non-chosen alternatives by using the hedonic modelling technique to obtain housing price data *for each dwelling*, and thus overcome the issue relating to the use of average housing prices. Housing price data are then used as input into a tenure and dwelling type choice model and a test undertaken to investigate the possible presence of price endogeneity. To this end, this paper uses an error component model where the error components are specified as a function of the residuals from the hedonic models of dwelling cost. This approach of testing and correcting for price endogeneity was first proposed by Blundell and Smith (1989) and recently applied to modelling residential location choice by de Palma et al. (2005). Technical specification is covered in the next section.

3. Methodology

3.1 The survey and data

An online survey was purposely designed to collect data for this project. A pilot survey was conducted on 25th July 2013 on a sample of 30 respondents in Sydney to test the comprehendability of the questionnaire and the workability of the database at the back-end. Minor edits were made as a result of the pilot survey and the main survey was conducted with the help of *pureprofile,* an online survey company, from 6th to 19th August 2013. A survey link was sent by *pureprofile* to its subscribed respondents living in the Sydney Greater Metropolitan Area (SGMA) including Sydney, Illawarra and Newcastle (see Figure 1). A preset sample of 2,000 valid respondents was contracted and sampling quotas were applied based on the proportion of dwellings in each residential postcode to the total dwellings in the SGMA. This aims to increase the geographical representativeness of the sample. The database was analysed on a daily basis when the survey was in progress to ensure that the geographical quotas were closely matched. A sample of 2,014 respondents was obtained with 19 respondents living outside the SGMA. Figure 1 shows the sample distribution by postcode of the place of residence. Postcodes with no respondents shown in Figure 1 are mostly bush areas.

The questionnaire consisted of four parts. The first part included questions relating to residential location, where respondents were asked to describe their current residential location (suburb and postcode), tenure type (own, rent, or other), dwelling type (detached, semi-detached including town house, or unit including apartment) and structural characteristics of the dwelling they live in, including plot size, last year of major renovation, number of bedrooms, bathrooms, separate toilets, living rooms, rumpus rooms, dining rooms, and garages. In this first part, the respondents were also asked to estimate the approximate value of their dwelling if it were to be sold or leased on the day they did the survey, regardless of whether they own or rent it. The second part included questions relating to the dwelling property such as building age, number of units within the block where the dwelling sits, frontage of the property, number of storeys, wall type, and the existence of a lift, garden, swimming pool, and balcony.

The third part of the questionnaire asked respondents to describe spatial characteristics of the dwelling in terms of the street environment (totally residential or mixed residential and commercial), and their evaluation (on a 10-point scale) of the levels of trees and vegetation, traffic, crime, air noise, on-street parking, and view of water surfaces (harbour, river, canal, lake). Respondents were also asked to estimate distance from their home to the closest primary and secondary schools (separately for private and public schools), university, corner shop, shopping centre, bus stop and train station as well as average travel times by car and public transport (PT) during the morning peak from home to the Sydney Central Business District (CBD). The main reasons for choosing the current place and a record of residential locations in the last 20 years were also asked to help choose location alternatives for choice set formation for the residential location modelling stage. The final part included a survey of household characteristics such as number of people, workers and non-workers living in the same dwelling, household income, household structure, and number of household cars.

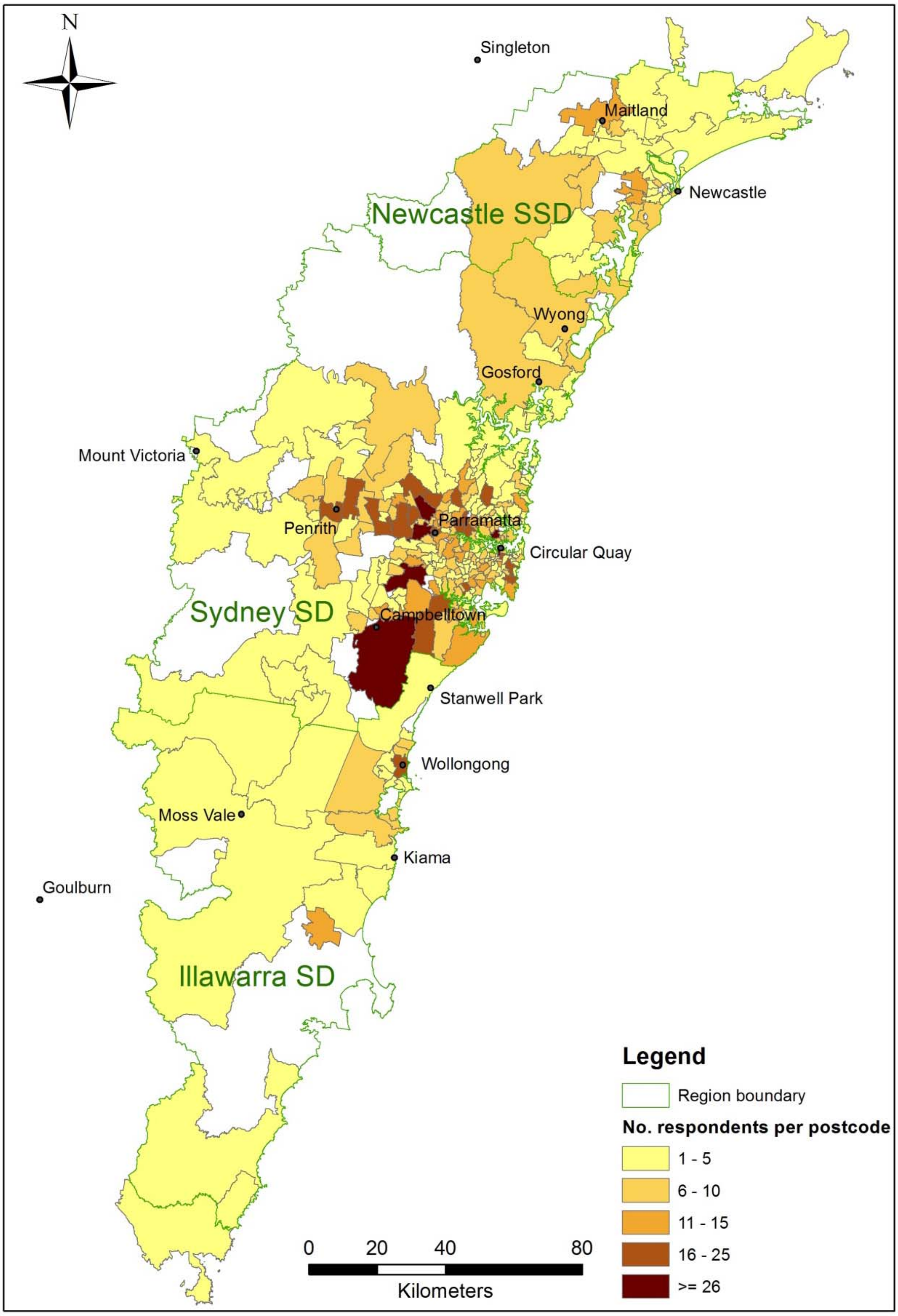

Figure 1: Spatial distribution of respondents to the residential location survey 2013

An extensive process of cleaning and validating the data reduced the sample to 1,557 usable dwellings from the initial 2,014 dwellings. Apart from 19 observations which were outside the study area, an additional 438 observations were removed based on one or multiple cleaning rules[1] which consider a realistic value range for each attribute and the consistency of information across different attributes. No effort was made to retain part of these observations as the remaining sample is sufficient for model estimation. Table 1 provides the distribution of dwellings by tenure type and dwelling type in the final sample, as compared to 2011 Census data for the SGMA. The final sample is very representative of the population in terms of dwelling type but less so for tenure type. That is, owners are somewhat over-represented in the sample (75%) as compared to the 2011 census (68%).

Table 1: Distribution of dwellings by tenure and dwelling type: surveyed sample vs. 2011 census

	2013 surveyed sample		**2011 Census**	
	Own	Rent	Own	Rent
Detached	862 (74%)	145 (37%)	911,940 (77%)	206,700 (38%)
Semi-detached	113 (10%)	66 (17%)	120,309 (10%)	90,388 (17%)
Unit	185 (16%)	186 (47%)	154,230 (13%)	249,003 (46%)
Total	1,160 (100%)	397 (100%)	1,186,479 (100%)	546,091 (100%)

Data sources: this study and 2011 Census.

3.2 Modelling approach

The analysis of housing choice requires a dwelling price for both the chosen and non-chosen alternatives, with the former being available in most revealed preference surveys. Supplementing price data for non-chosen alternatives is challenging given the potential issues of price endogeneity and heterogeneity discussed in section 2. In this paper, a hedonic modelling technique was employed to impute price data for non-chosen alternatives. Two separate hedonic models, one for owners and one for renters, were developed using data from the corresponding samples. With estimated values of both dwelling price and weekly renting cost being asked regardless of tenure type, we were able to identify the statistical residuals of the housing cost models that cannot be explained by the set of explanatory variables. The residuals represent the effect of all unobserved features that may partially be factored into the dwelling price. These were then incorporated into a choice model of dwelling and tenure type to test for the existence of price endogeneity.

A heteroscedastic error component model (ECM) was used to detect the existence of price endogeneity where heteroscedasticity is specified as a function of residuals from the hedonic models of dwelling cost (for example applications of ECM, see Brownstone and Train, 1998; Greene and Hensher, 2007). If the parameters associated with the residuals are not statistically significant, we cannot reject the non-hypothesis of price exogeneity. In other words, it is confident to conclude that the estimated effect of dwelling price on housing choice is not biased due to price endogeneity. The next section presents the estimation results.

[1] The rules are available on request.

4. Estimation results

4.1 Hedonic models of dwelling price and renting cost

The aim of a hedonic model is to find a broad set of explanatory variables that significantly explains variations in housing cost with a moderate level of multicollinearity (Löchl and Axhausen, 2010). In this paper, two sources of variations in housing cost are distinguished. The first is the structural characteristics of the dwelling such as the number of bedrooms, bathrooms, garages, and other amenities. The second source of variation is the difference in spatial setting characterised by access to local and regional facilities (e.g., school, university, public transport services, CBD, job centres), population density, average household income, and geographical region where the dwelling locates.

Table 2 presents the effect of structural and spatial characteristics on housing costs, measured as the natural log of dwelling price (for owning) and weekly rent (for renting). As indicated by the overall goodness of fit (R^2), the set of variables explains about 76% of the variation in dwelling price and about 66% of the variation in weekly renting cost. This is impressive for highly disaggregated data that preserves all the variance at the dwelling level. Statistically significant variables are essentially similar across two models with some variables such as land area, lift availability, and separate living room having a more specific effect on renting cost or dwelling price. Overall, all parameters have the expected sign and plausible magnitudes. Interpretation of the parameters shown in Table 2 is straightforward with the following highlighting the more important results.

Table 2: Hedonic models of dwelling price and weekly renting cost for SGMA 2013

	Y=ln(dwelling price). R^2 = 0.759			Y=ln (weekly rent); R^2 = 0.655		
Variable	**b**	**β**	**t**	**b**	**β**	**t**
Constant	12.1466	–	93.87	5.3794	–	29.87
Detached (1/0), base = unit	0.2576	.243	5.41	0.0721	.093	1.83
Semi-detached (1/0), base = unit	0.1989	.127	5.00	0.0704	.070	1.79
ln of land area (m^2)	0.0103	.055	3.45			
Number of bedrooms	0.0959	.216	10.05	0.1080	.297	6.49
Number of combined bathrooms and toilets	0.0662	.126	7.25	0.0400	.072	2.01
Separate dining room (1/0)	-0.0356	-.038	-2.02	0.0478	.061	1.68
Separate rumpus room (1/0)	0.0555	.060	3.16	0.1276	.116	3.22
Separate living room (1/0)	0.0344	.032	1.81			
Number of garages with door	0.0484	.089	4.94	0.1148	.222	5.85
Dwelling within a block (1/0)	-0.1759	-.152	-4.31			
Number of levels within the building	0.0185	.109	4.26			
Number of units within the building	-0.0010	-.073	-3.56			
Double brick wall (1/0), base = wood or other	0.1045	.111	4.49			
Veneer brick wall (1/0), base = wood or other	0.0550	.059	2.40			
Balcony availability (1/0)	0.0508	.055	3.18			
Lift availability (1/0)				0.1597	.150	4.08
Last major renovation (year)				-0.0011	-.038	-1.12
ln of building age (year)	-0.2873	-.561	-8.24	-0.1022	-.246	-1.97
ln of building age * ln of building age	0.0462	.529	7.46	0.0135	.183	1.42
Level of trees (1 = none, 10 = extensive)	0.0174	.078	4.85			
Level of crime (1= none, 10 = serious)	-0.0132	-.055	-3.44	-0.0055	-.033	-0.96
View of water surface (1=none, 10=very good)	0.0232	.121	7.17	0.0141	.093	2.70
Distance to public primary school (km)	-0.0056	-.023	-1.27			
Distance to public high school (km)	0.0070	.056	2.96			
Distance to university (km)	-0.0016	-.045	-2.49			
Sydney City (1/0), base = Illawarra	0.6444	.292	9.26	0.4723	.397	6.84
East (1/0), base = Illawarra	0.7815	.327	14.46	0.5296	.363	7.46
South (1/0), base = Illawarra	0.4359	.329	10.18	0.3532	.285	5.57
Inner West (1/0), base = Illawarra	0.5635	.231	10.21	0.4102	.240	5.27
Inner North (1/0), base = Illawarra	0.5476	.330	11.11	0.4081	.325	5.69
North (1/0), base = Illawarra	0.3149	.200	6.47	0.2588	.097	2.47
North East (1/0), base = Illawarra	0.6229	.246	11.32	0.4294	.180	4.24
Central West (1/0), base = Illawarra	0.2710	.232	6.62	0.2442	.245	4.16
North West (1/0), base = Illawarra	0.0104	.007	0.24	0.0755	.062	1.19
South West (1/0), base = Illawarra	0.0255	.014	0.57	-0.1080	-.073	-1.59
Central Coast (1/0), base = Illawarra	0.1323	.037	2.01	0.0617	.020	0.57
Hunter (1/0), base = Illawarra	-0.0257	-.014	-0.61	0.0043	.003	0.07
Median household income ('000$)	0.0003	.241	11.29	0.0001	.161	3.68
ln of population density ('000/km^2)	0.0140	.036	1.32	-0.0057	-.020	-0.34
Median journey-to-work PT travel time (min)	-0.0013	-.101	-3.45	-0.0014	-.149	-2.50

Of the structural characteristics significantly entering the models, dwelling type, the number of bedrooms, bathrooms, garages and building age appear to have the strong effect on dwelling price and renting cost (see standardised coefficients β). However, dwelling type has a more important role in explaining variations of dwelling price than weekly renting cost. This result is expected as owners' preference for a detached and semi-detached dwelling in contrast to a unit (the base) may be much stronger than renters' preference, given a typically longer term commitment of owners. The results suggest that, *ceteris paribus,* owners pay 25.8% and 19.9% more to own a detached and a semi-detached dwelling, compared to what they pay for a unit. In contrast, the difference in weekly renting cost is only about 7.2% for a detached dwelling and 7.0% for a semi-detached dwelling (see coefficients *b*). In addition, the effect of building age is not linear, with the quadratic terms of the building age variable entering as a statistically significant influence on the dwelling price model. This reflects the heritage of old dwellings in the study area, such as the Rocks precinct in Central Sydney.

Of the various spatial characteristics tested with the model, the level of crime, view of water surface, median household income of the residential postcode, and regional dummies were found to be statistically significant influences on dwelling price and weekly renting cost. Dwellings in the Eastern suburbs are the most expensive, as expected, with owners paying 78% more while renters pay 52% more than they do for a similar dwelling in Illawarra (the most southerly location in the SGMA). Dwellings in higher income neighbourhoods with a shorter median PT travel time to work are also valued higher, both for owning and renting, while access to school and university only influences owning price. Interestingly, distance to a public high school, positively correlated while distances to a university and primary school, are inversely correlated with dwelling price. A possible explanation is that dwelling owners are more concerned with traffic conditions and parking associated with dropping off and picking up high-school-aged students. Drop-offs and pick-ups also happen around primary schools but to a much lesser extent with walking being the main mode for escorting primary-school children and thus the effect is not significant (Ho, 2013). Dwellings closer to a university have a higher opportunity to lease out and this may be attractive to property investors, leading to a higher price.

Given the plausible findings discussed above, two models shown in Table 2 were used to estimate price data for non-chosen alternatives. This was undertaken by turning on and off the dummy variables representing dwelling type (i.e., detached and semi-detached) while retaining the same values for other attributes, assuming that alternative dwelling types with similar spatial and structural characteristics are always available in the same metropolitan region. This assumption is not strict given a mix of dwelling types in each of the metropolitan regions. Residuals of dwelling price were then calculated as the difference between the respondent's estimated values and those generated by the hedonic models. Both dwelling prices and residuals were then used as input into a model representing the choice of tenure and dwelling type.

4.2 Discrete choice models of tenure and dwelling type

Different choice models with varying substitution patterns are assumed and estimated using Nlogit 5. Table 3 presents the estimation results of the final ECM and NL models after exploring a number of alternative tree structures. The former, whose results are presented the first two columns, is specifically used to test for the existence of price endogeneity through the four error components. The first and second error components are the standard deviation of *iid* deviates that enter the utility for each dwelling type of a given tenure type. Following Brownstone and Train (1998), these error components are motivated by previous studies modelling dwelling type conditioned on tenure type (Boehm, 1982; Skaburskis, 1999). The third and fourth error components relate to the residuals of dwelling price and weekly renting cost models. These are used to detect price endogeneity.

Table 3: Estimation results of tenure and dwelling type choice model for SGMA 2013

	Error Components		Nested Logit	
Variable	**Coeff.**	**t-value**	**Coeff.**	**t-value**
Household income, owning alternatives	0.006	4.03	0.005	4.31
Family household, owning alternatives	0.782	3.47	0.933	5.45
Constant, owning alternatives	-2.431	-0.93	-2.974	-1.81
ln(dwelling price), generic for owners	-1.526	-6.89	-1.131	-8.43
Household size, own detached	0.709	8.78	0.320	6.99
Household size, own semi-detached	0.357	3.25	0.165	3.41
Distance to CBD, own detached	0.039	15.00	0.025	10.64
Distance to CBD, own semi-detached	0.016	3.45	0.015	6.82
Couple household, own detached	0.540	2.35	0.274	2.51
Couple household, own unit	0.194	0.73	0.099	0.83
Constant, own detached	-1.410	-4.73	-0.740	-5.17
Constant, own semi-detached	-1.441	-3.56	-0.663	-3.72
ln(weekly renting cost), generic for renters	-3.589	-11.4	-2.949	-11.58
Household size, rent detached	0.965	9.52	0.696	8.23
Household size, rent semi-detached	0.572	4.75	0.362	3.31
Distance to CBD, rent detached	0.041	14.44	0.028	9.97
Distance to CBD, rent semi-detached	0.031	8.90	0.020	5.69
Couple household, rent detached	0.027	0.08	-0.183	-0.63
Couple household, rent unit	-0.031	-0.12	-0.122	-0.59
Single parent HH, rent detached	0.579	1.80	0.528	1.73
Constant, rent detached	-4.124	-10.29	-3.144	-8.56
Constant, rent semi-detached	-3.153	-7.38	-2.469	-6.29
Logsum parameters				
Own			0.373	9.16
Rent			1.000	fixed
Error components				
Own	0.662	1.44		
Rent	0.030	0.09		
Residual of dwelling price model	2.623	1.70		
Residual of renting cost model	-4.874	-0.47		
Log likelihood at convergence	-1796.72			-1772.26
Log likelihood at zero	-2789.77			-2789.77
Log likelihood at market share	-2148.18			-2148.18
Sample size	1557			1557
Number of parameters	26			23
McFadden adjusted R^2 (vs. zero)	0.356			0.356
McFadden adjusted R^2 (vs. constant)	0.152			0.164

As can be seen in Table 3, two parameters associated with the residuals of dwelling price and weekly renting cost models are not statistically significant at the 5% level. This suggests that dwelling price is not endogenous with respect to the choice of tenure and dwelling type and the effect of dwelling prices on dwelling choice is consistently estimated. The focus is then on finding a model which gives the better overall model fit and provides behaviourally meaningful results.

The same explanatory variables are specified with nested logit (NL) models with two levels. A NL model which groups dwelling types under tenure type produces results which are behaviourally meaningful and consistent with utility maximisation. The estimation results and summary statistics of this NL model are presented in the last two columns of Table 3. Both ECM and NL models produce a behaviourally meaningful and consistent sign for estimated parameters but the latter is statistically superior to the former in terms of the overall model fit (McFadden adjusted R^2). Therefore, interpretations of estimated parameters and price elasticities are based on the NL model.

Household income and household structure significantly influence household choice of tenure as shown in the first two rows of Table 3. *Ceteris paribus*, family households are more likely to own than to rent a dwelling while differences in preference for tenure type are not statistically different for other household types. The positive parameter estimate associated with family households may reflect the strong effect of lifecycle on owner-occupation but this can also means these households have enough financial budgets to enter the property market. The effect of economic constraints or financial capability on dwelling ownership is reinforced by the positive parameter of household income, suggesting that high income households are more likely to be an owner than a renter.

Given tenure type, household choices of dwelling type are significantly correlated with housing costs, with both renters and owners preferring a less expensive dwelling of any type. However, larger households prefer a detached and a semi-detached dwelling to a unit (the base), presumably to have room for all household members. In order to so, larger households are found to trade-off the number of rooms with access to the CBD, as reflected by the positive parameters associated with distance to the CBD, both for renters and owners of detached and semi-detached dwellings.

4.3 Price elasticities of demand

A common feature of discrete choice models is that the magnitudes of estimated coefficients are not readily interpretable (Hensher et al., 2005). Therefore, elasticities of demand are derived and presented in Table 1. Dwelling prices and household income are selected for this analysis. Arc elasticities and the probability weighted sample enumeration (PWSE) aggregation method are used in reference to point elasticities and ‘sample average’ and ‘naive pooling’ aggregation methods (for in-depth discussion, see Hensher et al., 2005). The estimated coefficients and input data are used to replicate the final NL model in Excel which is used for simulation with a 10% increase in the interested variables. There are two reasons for doing so. First, Nlogit 5 does not simulate changes to attribute specified at the upper level of a NL model (i.e., household income is specified at the tenure choice level). Second, the natural log transformation is used for dwelling prices for model specification while price elasticities are desired for interpretation.

Table 1: Mean direct and cross arc elasticities of demand for tenure and dwelling type choice probabilities

	Own Detached	Own Semi	Own Unit	Rent Detached	Rent Semi	Rent Unit
Detached price	-0.80	1.69	1.24	0.60	0.54	0.40
Semi price	0.20	-2.66	0.32	0.05	0.06	0.08
Unit price	0.27	0.57	-2.01	0.06	0.11	0.19
Detached renting cost	0.24	0.17	0.12	-2.42	0.32	0.20
Semi renting cost	0.10	0.10	0.09	0.14	-2.76	0.14
Unit renting cost	0.22	0.34	0.48	0.27	0.40	-2.29
Household income	0.11	0.12	0.14	-0.34	-0.35	-0.34

Direct elasticities are in shaded cells.

Table 1 shows that the demand for all tenure and dwelling types is relatively elastic with respect to price of the same alternative (i.e., direct elasticities), except for the demand for owning a detached dwelling, where a 1% increase in its price results in a 0.80% decrease in the probability of owning it. In contrast, a 1% increase in the price of semi-detached dwellings will result in a 2.66% decrease in the choice probability of owning a semi-detached dwelling, *ceteris paribus*. The cross-elasticities exhibit a range from relatively elastic (e.g., between owner occupation of semi-detached and unit dwellings and detached price) to relatively inelastic between owner occupation and renting for all dwelling types as well as between owner-occupation of semi-detached dwelling and unit prices. This suggests that a given percentage increase in the price of a detached dwelling leads to a higher the percentage increase in the probability of home buyers choosing a semi-detached or a unit dwelling. However, the reverse is not observed in the data, with a given percentage change to the price of semi-detached and unit dwellings resulting in a smaller (marginal) percentage change in the choice probability of owning a detached dwelling.

With respect to household income, the demand for ownership increases as household income increases. However, income elasticities of demand for owning and renting any dwelling type are relatively inelastic, with a 10% increase in household income resulting in about 1.2% increase in owning and about 3.4% decrease in renting. The estimated income elasticities of demand for owning a dwelling in Sydney are at the low end compared to the range of (0.4 – 1.5) in the US and (0.1 – 1.0) in the UK markets (Ermisch, 1996; Rapaport, 1997). The differences in income elasticities of demand for housing across countries are likely due to different housing markets and the way in which household income is measured in this paper with single-year income tending to underestimate income elasticities of demand (Barrios García and Rodríguez Hernández, 2008).

5. Placement of the model in a general transport-location strategic framework

The proposed model for the choice of tenure and dwelling type is a part of a new version of TRESIS being currently developed for Sydney GMA by the Institute of Transport and Logistics Studies (ITLS) at the University of Sydney. At the time of writing, the model has not been applied, but will be incorporated into the framework of the new TRESIS system. The proposed model evolves from earlier versions by including tenure choice and considering a finer spatial resolution.

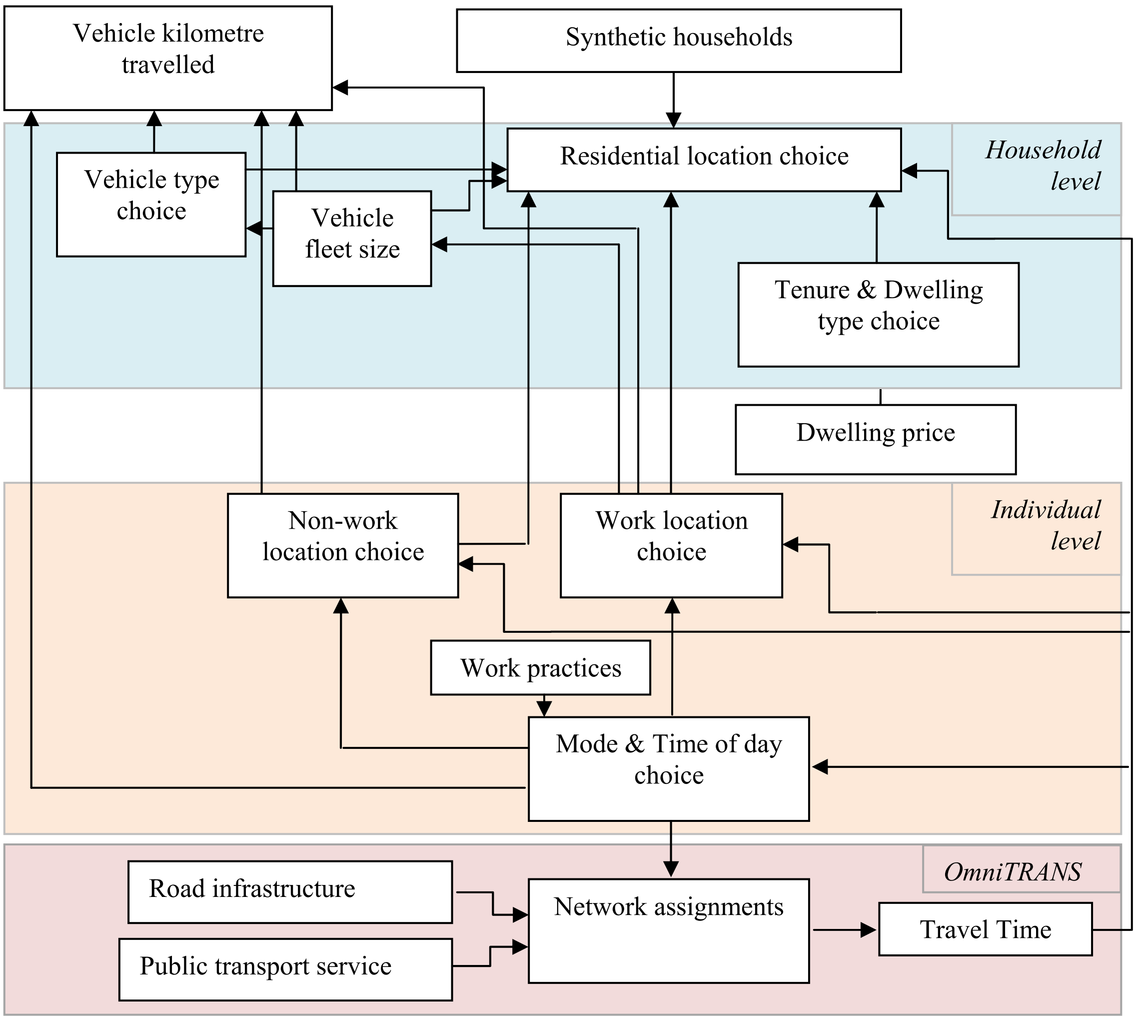

Figure 2: Overall framework of TRESIS passenger travel demand models

An overall structure of the under-development version of TRESIS passenger travel and location model system is presented in Figure 2. This includes a series of passenger travel demand models applied sequentially, either at a household level or an individual level, with feedbacks and links between modules. Apart from allowing for a flexible number of zones and a finer spatial resolution, this new version of TRESIS has several additional advanced features compared to the previous version described in Hensher (2002) and Hensher and Ton (2002). The first advancement is the introduction of a non-work location choice model, which is applied at the individual level and conditioned on household choice of residential location. The second advancement over the earlier versions is the joint choice model of travel mode and time of day. Data collection for these two models is in progress with the results to be reported in the near future.

As for the model developed in this paper, it is applied conditional on the choice of residential location. Once estimated and then calibrated, the developed tenure and dwelling type choice model is used to calculate the inclusive value (IV, or logsum) which differs across residential locations due to distance to the CBD, forming a profile of expected utility for the choice of tenure and dwelling type. The established IV will then be fed into residential location choice model, together with IVs from other models such as work and non-work location choices.

At the bottom of the passenger travel demand model system structure is the mode and time of day choice, which is treated as a joint decision. Output from this model together with road network and public transport services are fed into OmniTRANS for network assignment, allowing for re-calculation of travel times and skim matrices. New travel times are then updated to the choice system in an iterative manner until convergence in all levels is obtained.[2] The final results can then be used for policy evaluation including those relating to vehicle kilometre travelled, greenhouse gas emissions and many other spatial and aspatial policies (see for example, Hensher, 2002, 2008).

6. Discussion and conclusions

A major task involved in setting up a strategic travel demand model system such as TRESIS relates to data collection and imputation. This is particularly true when it comes to estimating residential and dwelling type choices which require both household information and dwelling characteristics, while most available data possess only one part but not both. In addition, for model development and application, dwelling price data are required for both chosen and non-chosen alternatives while revealed preference data only have price of the chosen alternative. This paper illustrated one way in which dwelling price data may be supplemented by the use of a hedonic modelling technique and an online survey of a representative sample. The imputed dwelling prices were then used to model the choice of tenure and dwelling type with price endogeneity being tested and shown not to exist.

Insight into drivers of dwelling price and renting cost in Sydney as well as its residents' choice of tenure and dwelling type was gained. Structural features of a dwelling appear to be the main drivers of dwelling price but the spatial setting also plays an important role, especially for owning. Generally, access to local and regional facilities added value to the dwelling but there exists a clear distinction between owners and renters in the valuation of these facilities. For example, the greater the amount of trees and vegetation in the local environment, the higher the value of the dwelling in the eyes of the owners, but this is not reflected as a value-added factor in the valuation of the renters when it comes to paying rent. Similarly, access to a university significantly increases the dwelling price but not the renting cost. However, being closer to a university may increase the chance a dwelling being leased out, which indirectly increases long-term investment return. Understanding these differences has important implications for property investment.

Given the dwelling prices, the choice of tenure type is mainly driven by a financial budget to enter the property market, with high income households and family households being more likely to be an owner. Derived price elasticities of demand for different dwelling types also reflect the financial barrier to property ownership in Sydney. An increase in the price of detached dwellings results in a strong switch to owning a semi-detached dwelling or a unit; however, an increase in the price of semi-detached or unit dwellings does not result in a significantly higher demand for detached dwellings. That is, a detached dwelling may be unaffordable to most home buyers due to the financial barrier, which explains why a reduced financial gap of owning a detached and a semi-detached/unit dwelling (due to an increased price of a semi-detached/unit dwelling or a reduced price of a detached dwelling) does not make

[2] Equilibrium assignment also occurs in the automobile market using vehicle prices and disequilibrium in the dwelling market using dwelling prices.

detached dwellings more attractive (i.e., cross and direct elasticities of detached dwellings are smaller than one).

This paper is part of a larger and ongoing project aiming to improve the current TRESIS system to allow for more powerful scenario analysis and evaluation of transport policy. For a complete TRESIS model system, it is necessary to link the choice model of tenure and dwelling type described in this paper with the residential location choice model.

References

Barrios García, J. A. & Rodríguez Hernández, J. E. 2008. Housing demand in Spain according to dwelling type: Microeconometric evidence. *Regional Science and Urban Economics,* 38**,** 363-377.

Blundell, R. W. & Smith, R. J. 1989. Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models. *The Review of Economic Studies,* 56**,** 37-57.

Boehm, T. P. 1982. A hierarchical model of housing choice. *Urban Studies,* 19**,** 17-31.

Brownstone, D. & Train, K. 1998. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics,* 89**,** 109-129.

Cho, C.-J. 1997. Joint choice of tenure and dwelling type: a multinomial logit analysis for the city of Chongju. *Urban Studies,* 34**,** 1459-1473.

de Palma, A., Motamedi, K., Picard, N. & Waddell, P. 2005. A model of residential location choice with endogenous housing prices and traffic for the Paris region.

Ermisch, J. 1996. The Demand for Housing in Britain and Population Ageing: Microeconometric Evidence. *Economica,* 63**,** 383-404.

Greene, W. H. & Hensher, D. A. 2007. Heteroscedastic control for random coefficients and error components in mixed logit. *Transportation Research Part E: Logistics and Transportation Review,* 43**,** 610-623.

Hensher, D. A. 2002. A systematic assessment of the environmental impacts of transport policy. *Environmental and Resource Economics,* 22**,** 185-217.

Hensher, D. A. 2008. Climate change, enhanced greenhouse gas emissions and passenger transport–What can we do to make a difference? *Transportation Research Part D: Transport and Environment,* 13**,** 95-111.

Hensher, D. A., Rose, J. M. & Greene, W. H. 2005. *Applied choice analysis: a primer*, Cambridge University Press.

Hensher, D. A. & Ton, T. 2002. TRESIS: A transportation, land use and environmental strategy impact simulator for urban areas. *Transportation,* 29**,** 439-457.

Ho, C. 2013. *An investigation of intra-household interactions in travel mode choice.* PhD Thesis, The University of Sydney.

Löchl, M. & Axhausen, K. W. 2010. Modelling hedonic residential rents for land use and transport simulation while considering spatial effects. *Journal of Transport and Land Use,* 3.

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

Skaburskis, A. 1999. Modelling the choice of tenure and building type. *Urban Studies,* 36**,** 2199-2215.

Tu, Y. & Goldfinch, J. 1996. A two-stage housing choice forecasting model. *Urban Studies,* 33**,** 517-537.

Yates, J. & Mackay, D. F. 2006. Discrete choice modelling of urban housing markets: A critical review and an application. *Urban studies,* 43**,** 559-581.