AUSTRALIAN
SCHOOL OF BUSINESS ™
THE UNIVERSITY OF NEW SOUTH WALES

www.business.unsw.edu.au

School of Economics
Australian School of Business
UNSW Sydney NSW 2052 Australia

http://www.economics.unsw.edu.au

Flexible Spatial and Temporal Hedonic Price Indexes for Housing in the Presence of Missing Data

Iqbal Syed, Robert J. Hill and Daniel Melser

School of Economics Discussion Paper: 2008/14

ISSN 1323-8949
ISBN 978-0-7334-2700-8

Flexible Spatial and Temporal Hedonic Price Indexes for Housing in the Presence of Missing Data

Iqbal Syed[a], Robert J. Hill[b,a] and Daniel Melser[c]

Draft: 26 September 2008

We propose a flexible hedonic methodology for computing house price indexes that uses multiple imputation (MI) to account for missing data (a huge problem in housing data sets). Ours is the first study to use MI in this context. We also allow for spatial correlation, include interaction terms between characteristics, between regions and periods, and between regions and characteristics, and break the regressions up into overlapping blocks of five consecutive periods (quarters in our case). These features ensure that the shadow prices are flexible both across regions and time. This flexible structure makes the derivation of price indexes from the estimated regression equations far from straightforward. We develop innovative methods for resolving this problem and for splicing the overlapping blocks together to generate the overall panel results. We then use our methodology to construct temporal and spatial price indexes for 15 regions in Sydney, Australia on a quarterly basis from 2001 to 2006 and combine them to obtain an overall price index for Sydney. Our hedonic indexes differ quite significantly from the official index for Sydney published by the Australian Bureau of Statistics. We also find clear evidence of convergence in prices across regions from 2001-3 (while prices were rising), and divergence thereafter. We conclude by exploring some of the implications of these empirical findings. (**JEL.** C43, E01, E31, R31)

Keywords: Real estate; House prices; Hedonic price index; Missing data; Multiple imputation; Spatial correlation

[a]School of Economics, University of New South Wales, Sydney 2052, Australia: i.syed@unsw.edu.au, r.hill@unsw.edu.au

[b]Department of Economics, University of Graz, Universitätsstrasse 15/F4, 8010 Graz, Austria: robert.hill@uni-graz.at

[c]Moody's Economy.Com, Level 10, 1 O'Connell Street, Sydney, NSW 2000, Australia: DMelser@economy.com

1 Introduction

Movements of house prices are of great interest to households, investors in the housing market, and policy makers pursuing diverse goals such as monitoring and tracking underlying inflation and maintaining social equity. Housing accounts for more than 50 percent of the private capital stock and 30 percent of household expenditure in the US (see Englund, Quigley and Redfearn 1998) and around 60 percent of household assets in Australia (see Hansen 2006). It is not surprising therefore that movements in house prices can significantly impact on the economy. Indeed, Case, Quigley and Shiller (2005) find that changes in house prices have a greater impact than changes in stock market prices on household consumption. The importance of the housing market has been clearly demonstrated by the recent economic turmoil in various countries (e.g., the US, UK, Spain and Ireland) triggered by the subprime mortgage crisis in the US.

It is important, therefore, that movements in house prices and regional differences are accurately measured. This is difficult since every house is different. Ideally what is required is a quality adjusted index that compares like with like from one period (or region) to the next. The most common approach taken by national statistical agencies is to track the average or median price of houses sold. Such an index fails to make a quality adjustment, and hence may be misleading when the mix of houses sold changes over time or space, as may well be the case in times of stress when accurate measurement is most urgently required.

One way to control for quality change is to use hedonic methods. Hedonic models regress the price of a product on a vector of characteristics. Most applications of hedonics focus on goods subject to rapid technological change, particularly computers (see for example Berndt, Griliches and Rappaport 1995, Pakes 2003, Triplett 2004, and Diewert, Heravi and Silver 2007). In a housing context, the characteristics can include physical factors such as land area, the number of bedrooms and bathrooms, and locational factors such as longitude, latitude, distance to the nearest shopping center, etc. Within a regression framework, 'pure price' changes from one period to the next or one region to another can be captured after controlling for qualitative and compositional changes. The implementation of hedonic regression methods requires detailed information on the physical and locational attributes of the sales records. The more information provided for each house sale, the better is the chance of capturing the 'pure price' change.

In this study, we propose a methodology for constructing hedonic price indexes for housing on panel data sets that is very flexible. Our first task is to address the fundamental problem of missing data. In our data set around 59 percent of sales observations are missing one or more of the core characteristics. Typically, observations with missing information are simply deleted (the default setting in most statistical packages). List-wise deletion, however, can impart bias to the standard errors and the price indexes themselves. We address this problem using multiple-imputation techniques developed by Rubin (1987) to fill in the gaps in the data set, prior to estimating the hedonic model. Our study is the first to apply these state-of-the-art methods in a housing context. This in itself is an important contribution, given the pervasiveness of the missing-data problem in housing data sets.

Location is an important factor in the housing market. It is necessary therefore that the hedonic model takes account of spatial correlation in the price data. We do this using methods developed by Anselin (1988). We also correct for heteroscedasticity by using the fact that the errors can be expressed as a function of some of the explanatory variables.

Flexibility is achieved through the inclusion of interaction terms between characteristics, between regions and periods, and between regions and characteristics. The interaction terms are statistically significant in almost all periods. The inclusion of all these interaction terms ensures that there is no substitution bias in the resulting price indexes.[1] In addition, the regressions are broken up into overlapping blocks of five consecutive periods (quarters in our case). This allows the shadow prices to adjust over time, and ensures that the results for earlier years do not change when a new year is

[1]Hill and Melser (2008b) show how fixed shadow prices can generate substitution bias.

added to the data set. This criterion – sometimes referred to as temporal fixity (see Hill 2004) – is particularly important for national statistical offices and other organizations that are required to compute indexes on an ongoing basis.

This flexible structure, however, makes the derivation of price indexes from the estimated regression equations far from straightforward. The presence of so many interaction terms implies that innovative methods are required to derive the price indexes for each overlapping block. Further innovations are required to splice these overlapping blocks together to generate the overall panel results. Our preferred method of splicing does this without distorting any of the temporal indexes (which are generally considered more reliable since they tend to be less affected by omitted variables bias).

We use our methodology to construct temporal and spatial price indexes for 15 regions in Sydney, Australia on a quarterly basis from 2001 to 2006 from a data set consisting of 418,877 sales observations. We also consider how best to combine the 15 regional indexes to obtain an overall price index for Sydney.

Our empirical results raise some interesting issues. In particular, our hedonic indexes suggest that the rise in house prices from 2001 to 2003 was not as large as indicated by the official index published by the Australian Bureau of Statistics (ABS). Conversely, we find that the fall in house prices since 2003 has been rather smaller than indicated by the ABS index. The fact that the Sydney housing market peaked in 2003 also allows us to compare regional pricing patterns in a rising and falling market. We find clear evidence of convergence in prices across regions while house prices were rising, and divergence thereafter.

2 The Missing Data Problem

The data set used in this study contains information on the sales of dwellings across 191 postcodes in Sydney from the beginning of 2001 to the end of 2006, a period of 24 quarters. The data was purchased from a private housing data provider, Australian Property Monitors, which obtains some of its information from the Office of the New

South Wales Valuer-General and supplements this with additional data (such as missing information on the number of bedrooms and bathrooms). We have supplemented further by purchasing information on coordinates (latitude and longitude) for the location of each dwelling. The data set used contains around 418,877 sales records.[2] Some summary information on the number of observations and prices of dwellings are provided in Table 1. Houses and units account for 53 and 47 per cent of the observations, respectively.

Insert Table 1 Here

The available information on physical attributes are *type of dwellings*, *bedroom count*, *bathroom count* and *lot size*. The *date of contract* was recorded for each dwelling, from which the *quarter* in which the sale took place was obtained. Detailed addresses were available, which enabled us to identify locations of the dwellings on a Cartesian space represented by longitudes and latitudes. In addition, the geographical area is divided into 191 *postcodes*.

Insert Figure 1 Here

We use the postcodes to divide Sydney into 15 regions. These regions are the same as those used by Residex, a private housing data provider, and also accord with our own idea of Sydney housing sub-markets.[3] The most number of sales in our data set

[2]The original data set had more observations. Some observations are excluded because they are considered to be 'outliers' or 'extreme observations'. For example, a *1 bed-1 bath* small unit in the suburb 'Neutral Bay' was recorded to be sold for 32,750,000 dollars in 2001, where the median price for that suburb is 520,000 dollars. There are some extreme numbers, at both ends, with respect to physical attributes such as bedroom and bathroom counts and lot size. Faced with the practical difficulty in identifying 'outliers', especially in a multi-dimensional context, we have decided to exclude 1 per cent of the observations from both tails of the distribution of dwelling prices and lot size. Figure 1 provides the distribution of prices and natural logarithm of prices before and after exclusion of extreme observations.

[3]Residex considers 16 regions. However, we do not have sales records for Residex's Campbelltown region. The Residex regions used, with postcode ranges in brackets, are: A=Inner Sydney (2000 to 2020), B=Eastern Suburbs (2021 to 2036), C=Inner West (2037 to 2059), D=Lower North Shore (2060 to 2069), E=Upper North Shore (2070 to 2087), F=Mosman-Cremorne (2088 to 2091), G=Manly-

were recorded in the region *Fairfield-Liverpool* (42,564), whereas the least was recorded in the region *Mosman-Cremorne* (10,848).

A large number of sales observations are missing information on core characteristics – bedroom or bathroom counts or lot size. This is particularly a problem for the present analysis because the hedonic approach is data-intensive and requires detailed characteristics information in order to implement. The subset of data with information on all the characteristics – bedroom and bathroom counts and lot size – consisted of 172,627 observations. Though this is still a very large data set, it is a reduction of around 59 per cent. The exclusion of such a large portion of the data would result in a loss of efficiency and, additionally, might lead to biased estimates.

The critical question that needs to be addressed is why some of the characteristics data are missing. The reason for *missingness* determines the way it impacts the estimates and provides guidance to the types of statistical procedures that are needed to arrive at estimates possessing desirable statistical properties. For example, while list-wise deletion might be an efficient way of dealing with missing data under certain situations, in other situations it might produce estimates that are biased, inefficient and unreliable.

In modern missing data procedures, missingness is considered to be a probabilistic phenomenon (see Schafer and Graham, 2002). Since in most cases missingness is beyond the researcher's control, we might not know its specific distribution. The way to proceed is to make assumptions on the randomness and how the missingness is related to the values of the missing items themselves. Distributions of missingness are classified according to the nature of their relationship with the values of the missing items. These assumptions are usually 'untestable'. However, investigation into the causes of missing

Warringah (2092 to 2109), H=North Western (2110 to 2126), I=Western Suburbs (2127 to 2145), J=Parramatta Hills (2146 to 2159), K=Fairfield-Liverpool (2160 to 2189), L=Canterbury-Bankstown (2190 to 2200), M=St George (2201 to 2223), N=Cronulla-Sutherland (2224 to 2249), Campbelltown (2552 to 2570), O=Penrith-Windsor (2740 to 2777). Henceforth we refer to these regions by their alphabetical prefixes.

data might provide guidance in making realistic assumptions.

Adopting a generic notation, let us denote the complete data set by $Y_{com} = (Y_{obs}, Y_{mis})$, where Y_{obs} and Y_{mis} refer to the observed and missing data, respectively. Let R denote the indicator for missingness, taking the value of 1 if a characteristic observation is missing and 0 otherwise. Rubin (1976, 1987, p.53) defines three distributions for missingness: missing at random (MAR), missing completely at random (MCAR) and missing not at random (MNAR). These are described in equations (1), (2) and (3), respectively:

$$P(R \mid Y_{com}) = P(R \mid Y_{obs}), \quad \text{(MAR)} \tag{1}$$

$$P(R \mid Y_{com}) = P(R), \quad \text{(MCAR)} \tag{2}$$

$$P(R \mid Y_{com}) = P(R \mid Y_{mis}). \quad \text{(MNAR)} \tag{3}$$

In a housing context, if the likelihood of the bedroom count being missing depends on the bedroom count, then the missingness is MNAR. If the likelihood of the bedroom count being missing depends not on its own value but on the value of other variables (e.g., the period of sale or the suburb) then the missingness is MAR. If the likelihood of the bedroom count being missing does not depend on any of the variables then the missingness is MCAR.

The MCAR assumption is the easiest to handle. Since the missing data values are simple random samples of all data values, the common procedure of list-wise deletion provides unbiased estimates.[4] However, standard errors are generally larger, resulting in inefficient estimates because in a reduced sample less information is utilized. If missingness has a MAR distribution, the reduced sample is not a simple random sample of the complete data set and, therefore, estimation might be biased.[5] For a similar

[4]This procedure is also known as 'case deletion' or 'complete case analysis' and is the default procedure for most statistical modelling in different statistical packages.

[5]Under MAR, missingness is related to the observed values of some other variables. Missingness is MCAR conditional on those variables. Therefore, the data set with complete information is not a simple random sample of the whole data set.

reason, though with greater force, list-wise deletion produces biased estimates under MNAR (see Allison 2002).

We assume here that the missing data have a MAR distribution. This is for a simple reason, which applies to many other data sets. The data set has become better over time because of an improved data collection process and by collectors being more careful about the quality of the data. This means that the probability of missingness depends on the 'time of sale' – an important determinant for the price of dwellings. Table 2 provides support for this hypothesis. While 69.6 percent of the sales records missed at least one characteristic in 2001, the percentage declined steadily over the years and reached 31.3 percent in 2006. This improvement can be seen individually for each of the core characteristics.

Insert Table 2 Here

Additionally, correspondence with the data provider revealed that characteristics information are available more for observations belonging to richer than poorer suburbs. Table 2 shows that regions with higher median dwelling prices (see Table 1) tend to have less missing data than regions with lower median prices. For instance, F is the most and O is the least expensive region, and the missing data in the sales records of these regions amounts to 40.8% and 68.4%, respectively. In order to check on this further, we identified five least and most expensive postcodes in terms of median dwelling prices. These are shown in Table 3. Table 3 shows that missingness varies substantially between these two sets of postcodes.

Insert Table 3 Here

As mentioned earlier, the assumption on the distribution of missingness is not directly testable since missing data are unrecoverable. However, as discussed above, Tables 1, 2, and 3 provide some evidence that missingness depends on some of the other variables in the data set and, therefore, that the MAR assumption may be appropriate.[6]

Little and Rubin (1987) argue that using explicit models is better than informal

[6] Collins, Schafer and Kam (2001) demonstrate that an erroneous MAR assumption often only has a minor impact on the estimates and inferences.

procedures such as editing and mean substitution in handling missing data. The current literature recommends two general approaches developed by Rubin (1976, 1987): maximum likelihood (ML) and Bayesian multiple imputation (MI).[7] Schafer and Graham (2002, p.173) in reviewing various old and new procedures for handling missing data summarized the situation as follows:

> Although other procedures are occasionally useful, we recommend that researchers apply likelihood-based procedures, where available, or the parametric MI methods described in this article, which are appropriate under general MAR conditions ... ML and MI under the MAR assumption represent the practical state of the art.

In this study, we follow the MI approach. In MI each missing value is filled in by $m > 1$ simulated values prior to statistical analysis. Since the early 1990s, the MI approach has gained prominence mostly because of rapid improvement in computing power enabling generation of thousands of simulated values in a short time, and the development of new Bayesian simulation methods in the late 1980s (see Schafer 1997). In the case of an arbitrary missing data pattern, a Markov Chain Monte Carlo (MCMC) approach that assumes multivariate normality can be used to impute missing values. Implementation involves two steps: the imputation I-step and posterior P-step. The two steps are iterated long enough so that the distribution of the simulated data converges to a stationary distribution. The I-step draws $Y_{mis}^{(t+1)}$ from $p(Y_{mis} \mid Y_{obs}, \theta^{(t)})$ where $\theta^{(t)}$ is the estimated parameter at the t^{th} iteration.[8] The P-step draws $\theta^{(t+1)}$ from

[7] The ML and MI approaches have the same optimal properties; they produce estimates that are consistent and asymptotically efficient as well as asymptotically normal when the data are MAR. The advantage of MI over ML is that it can be applied to virtually any kind of data and any kind of model (see Allison 2002). However, the MI approach also has its limitations – it produces different results each time it is applied to the data, though the difference is expected to be minor.

[8] One has the option of using informative or non-informative priors. In this study we have used non-informative priors. The initial estimates, $\theta^{(1)}$, are the means and covariance matrices obtained from Y_{obs}. Rubin (1996) recommends the use of as many variables as possible (which are expected to be correlated) when doing multiple imputation. We use dwelling price, period of sale, postcode address, dwelling type, bedroom and bathroom counts, lot size and the longitude and latitude of each observation.

$p(\theta \mid Y_{obs}, Y_{mis}^{(t+1)})$. The two steps are iterated n times creating a Markov chain

$$(Y_{mis}^{(1)}, \theta^{(1)}), (Y_{mis}^{(2)}, \theta^{(2)}), \ldots, (Y_{mis}^{(n)}, \theta^{(n)}).$$

The process is continued until convergence is attained at a stationary distribution $p(Y_{mis}, \theta \mid Y_{obs})$. Once the convergence is attained, we can simulate an approximately independent draw of the missing values. A detailed description of the method is available in Schafer (1997).[9] Repetition of the above process m times creates m imputed data sets.

The parameter estimates obtained from m imputed data sets can be combined using Rubin's (1987) methodology. Rather than combining the imputations for each missing observation prior to estimating the hedonic model, Rubin's approach requires us to estimate the hedonic model separately for each imputed data set and only then combine the results.

Let $\hat{Q}_j$ and $\hat{U}_j$ be the estimated regression coefficients and standard errors of the regression coefficients obtained from the j^{th} imputed data set. The overall estimate is the average of the m estimates:

$$\bar{Q} = m^{-1} \sum_{j=1}^{m} \hat{Q}^{(j)}.$$

The variance of $\bar{Q}$ is:

$$V = \bar{U} + (1 + \frac{1}{m})B,$$

where $\bar{U} = m^{-1} \sum_{j=1}^{m} U^{(j)}$ is the within imputation variance, m the number of imputed data sets, and $B = (m-1)^{-1} \sum_{j=1}^{m} (\hat{Q}^{(j)} - \bar{Q})^2$ the between imputations variance. For hypothesis testing, Rubin (1987) recommends the following statistic:

$$S = V^{-1/2}(\bar{Q} - Q),$$

[9]We implement MI using a *proc mi* routine in-built in the SAS 9.1 software. See Yuan (2000) for a detailed discussion of various options and how SAS implements MI. For other implementation options with step-by-step instructions, see the resource webpage maintained by J. W. Graham (http://methodology.psu.edu/resources.html).

where Q denotes the value of $\bar{Q}$ assumed under the null hypothesis. Rubin shows that S has a t distribution with $v = (m-1)\{1 + \bar{U}/[(1+m^{-1})B]\}^2$ degrees of freedom. The degrees of freedom varies from $m-1$ to infinity depending on the rate of missing information. When the degrees of freedom are large, the variance is precisely estimated and hence not much is gained by increasing the number of imputed data sets m.

3 The Estimated Hedonic Model

3.1 Specification of the hedonic model

The hedonic method dates back at least to Court (1939), and was revived by Griliches (1961). The conceptual basis of the approach was laid down by Lancaster (1966) and Rosen (1974). The two main approaches which have been used in practice are the time-dummy method and the hedonic imputation method (see ILO 2004 and Triplett 2004). Our focus here is on improving and extending the former method.[10]

For the purpose of illustration, we specify a relatively straightforward panel version of the time-dummy method. Here we pool across all the regions and periods in the sample and estimate the region-time specific fixed effects and shadow prices of characteristics.[11]

$$\ln(p_{kth}) = \alpha + \sum_{\tau=2}^{T}\sum_{\kappa=1}^{K} \delta_{\kappa\tau} b_{\kappa\tau h} + \sum_{c=2}^{C} \theta_c z_{ch} + \epsilon_{\kappa\tau h} \quad \text{for} \quad \begin{array}{l} h = 1, \ldots, H_{kt}, \\ k = 1, \ldots, K \\ t = 1, \ldots, T \end{array} \qquad (4)$$

In (4), $k = 1, \ldots, K$ are the regions, $t = 1, \ldots, T$ are the periods and $c = 1, \ldots, C$ are the characteristics. The dependent variable is the natural logarithm of the price

[10]The hedonic imputation method is discussed in detail in Hill and Melser (2008a, 2008b).

[11]This method was first proposed by Aizcorbe and Aten (2004), who refer to it as the time-interaction-country product dummy method. Hill and Melser (2008b) refer to it as the region-time dummy method.

of an observation belonging to region-period kt. The dummy variable b_{kth} takes the value 1 if the observation h is from region-period kt, and zero otherwise. z_{ch} denotes a characteristic or attribute of a dwelling. In a housing context, typically most of the characteristics take the form of dummy variables. The primary interest lies in the coefficients δ_{kt} which measure the region-period specific fixed effects on the logarithms of the price level after controlling for the effects of the differences in the attributes of the dwellings. The advantage of this region-time-dummy model is that the price index $P_{js,kt}$ between region-periods js and kt is derived directly from the δ_{kt} coefficients as follows:

$$\hat{P}_{js,kt} = \exp(\hat{\delta}_{kt} - \hat{\delta}_{js}). \tag{5}$$

In fact, it can be shown that this index is a biased estimate of the desired population parameter due to the fact that we are taking a nonlinear transformation of a random variable (see Garderen and Shah 2002). A better approach, following Kennedy's (1981) suggestion, is to use the adjusted index, $\tilde{P}_{js,kt}$, which will be approximately unbiased:

$$\tilde{P}_{js,kt} = \exp\left[\hat{\delta}_{kt} - \hat{\delta}_{js} - \frac{\mathrm{Var}(\hat{\delta}_{kt}) + \mathrm{Var}(\hat{\delta}_{js})}{2}\right]. \tag{6}$$

We find that the difference between $\hat{P}_{js,kt}$ and $\tilde{P}_{js,kt}$ in practice is small. The price indexes are the same up to four decimal places. Hence, in the more complicated models that follow, to simplify matters we do not make this correction.

As can be seen, the indexes $\hat{P}_{js,kt}$ are derived from simple transformations of the estimated region-period dummy variables. This simplicity comes at a price. The region-period dummies and dwelling characteristics enter the hedonic function additively. In other words, the function exerts a restriction on the potential interactions between region-periods and the characteristics set. For example, if bedroom counts is a characteristics then the shadow price of, say, two bedrooms is forced to be the same between two regions. In a temporal context it constrains the value of two bedrooms to remain the same over time. Without explicitly testing the significance in the difference in explanatory power, imposition of such restrictions seems unwarranted. In addition to the standard problems of misspecification, Hill and Melser (2008b) show how these restrictions can introduce systematic bias into the results.

Our objective here is to construct a flexible version of the region-time-dummy hedonic model that allows for interactions between characteristics, between characteristics and region-periods, and that allows the shadow prices of characteristics in each region to evolve over time. While the interaction between region-period dummies and characteristics complicates the derivation of indexes through the involvement of more parameters, we show how the region-period price indexes can still be obtained in a reasonably straightforward way.

Our generalized version of the region-time-dummy model takes the following form:

$$\ln(p_{kth}) = \alpha + \sum_{\tau=2}^{T} \beta_\tau q_{\tau h} + \sum_{\kappa=2}^{K} \gamma_\kappa r_{\kappa h} + \sum_{\kappa=2}^{K} \sum_{\tau=2}^{T} \delta_{\kappa\tau} b_{\kappa\tau h} + \sum_{m=2}^{M_\kappa} \eta_{\kappa m} d_{\kappa m h} + Z\Theta + u_{kth},$$

$$\text{for } h = 1, \ldots, H_{kt}$$

$$k = 1, \ldots, K,$$

$$t = 1, \ldots, T, \qquad (7)$$

where $q_{\tau h}$ are dummy variables such that $q_{\tau h}$=1 if the observation h is from period t and zero otherwise. Similarly, $r_{\kappa h}$=1 if the observation h is from region k and zero otherwise. The dummy variables $b_{\kappa\tau h}$ denote interactions between periods and regions taking the value of 1 if the observation h is from region-period kt and zero otherwise. The postcode dummies are denoted by $d_{\kappa m h}$, where $d_{\kappa m h}$=1 for observation h's postcode and zero otherwise.

Z is a set of quality characteristics. It includes the dwelling type, number of bedrooms, bathrooms, lot size, and two-way interactions among this set of attributes. In addition, each of these attributes is allowed to interact with regions. A detailed exposition of Z is provided in equation (8).

$$Z\Theta = \sum_{c_i=2}^{C_i} \theta_{c_i} z_{c_i h} + \sum_{c_i=2}^{C_i} \sum_{c_j=2_{i \neq j}}^{C_j} \eta_{c_i c_j} z_{c_i c_j h} + \sum_{\kappa=2}^{K} \sum_{c_i=1}^{C_i} \xi_{\kappa c_i} z_{\kappa c_i h} \quad \text{for } i, j = 1, \ldots, I. \qquad (8)$$

In (8), $i, j = 1, \ldots, I$ are the quality characteristics (bedrooms, bathrooms, lot size, dwelling type) and $c_i = 1, \ldots, C_i$ denote the attributes of the characteristic i (e.g., the number of bedrooms for the case of the bedrooms characteristic). The lot size and the square of lot size are included in the equation as continuous variables. The other characteristics take the form of dummy variables. The equation includes two-way interaction terms among the characteristic attributes, which are denoted by $z_{c_i c_j}$. In addition Z includes interactions between the regions and characteristics.

3.2 Accounting for spatial correlation

Basu and Thibodeau (1998) document two reasons why positive spatial correlation may exist. First, neighborhoods tend to develop at the same time resulting in dwellings having similar structural characteristics. Second, dwelling in a neighborhood share the same locational amenities. Many of the price determining factors shared by neighborhoods are difficult to document explicitly. However, the influence of these potentially 'omitted' variables are contained in the neighboring prices. Therefore, in the course of predicting house prices or undertaking regression analysis, one should work with a mechanism which takes account of this information.

In order to provide an illustration of the locational dependence of house prices in our data set, two graphs of dwelling prices have been plotted against distance to central business district (CBD) and a sea beach. The distances are the euclidian distances in R^2 calculated from the locations defined by longitudes and latitudes for each dwelling. The plots are provided in Figure 2.[12] Note that that none of the plots are flat, implying the importance of location in the determination of house prices. Note also that the dwelling prices are influenced by many other determinants including other locational factors and, since these are not controlled for in Figure 2, the curvature of the graphs should only be interpreted as indicative rather than capturing a proper causal effect.

[12] With regard to the first diagram, an arc was drawn from the CBD area (in postcode 2000) to the deep west of the Sydney metropolitan area. The dwellings of the postcodes through which the arc passed were included in drawing the graphs. These postcodes are 2000, 2007, 2008, 2042, 2048, 2049, 2203, 2131, 2191, 2192, 2190, 2199, 2143, 2162, 2163, 2160, 2161, 2165, 2164, 2766, 2760, 2770, 2747 and 2750. For the second diagram, four adjacent postcodes were considered: 2031, 2032, 2033 and 2034. The sea beach used for measurement is in postcode 2034.

Insert Figure 2 Here

The presence of spatial correlation implies that the Gauss-Markov assumptions are violated. At the very least, this implies inefficient estimators and incorrect standard errors, leading to biased inferences (see Anselin 1988 and Basu and Thibodeau 1998). More precisely, positive correlation will cause upward bias in the t-statistics. Its impact can potentially also cause biased in the estimators themselves. If the spatially correlated omitted variables are also correlated with the included variables, then the estimated coefficients can be biased even in large samples.[13] Therefore, we need to incorporate spatial dependence into the model.

To do this, we first need to find the neighbors of each observation. By using information on whether two spatial units have a common *border* or *edge*, one can identify neighboring spatial units. Anselin (1988) considers two spatial units as 'contiguous' if they have a common border of non-zero length. Neighboring observations can also be defined with respect to 'distance' between two observations. Under this scheme, it is possible to either keep the number of neighbors fixed, from 1 to n or to keep the distance between the neighbors fixed. Typically, a euclidian distance function is used to calculate the distance using locational information represented by latitudes and longitudes.

Second, we need to map the calculated distances to spatial dependence. This can be achieved in various ways. A common feature of these functions is that the closer are two observations, the higher is the strength of the spatial dependence.

In order to provide a structure to the assumed spatial relationship, a matrix commonly known as a *spatial weight* or *contiguity* matrix is constructed. The nature of the spatial dependence is specified in the spatial weight matrix. There are a large number of ways to construct the matrix. Each cell in the matrix can be the inverse distance or a function of the inverse distance between neighboring observations. Alternatively,

[13]The existence of spatial correlation in the residuals of the traditional hedonic equations have been reported in many empirical studies. For example, see Can (1990), Basu and Thibodeau (1998), Bourassa, Hoesli and Peng (2003), Pace and Gilley (1997), and Pace, Barry, Clapp and Rodriguez (1998).

it is possible to construct a matrix of ones and zeros, with 'ones' denoting neighboring observations and 'zeros' otherwise.[14] An elaborate discussion of alternative ways of constructing a spatial weights matrix is provided in Kelejian and Robinson (1995).

For a data set of n observations, the matrix has the dimension $n \times n$. The ith row of the matrix specifies the spatial dependence of the ith observation with the other $n - 1$ observations. The elements of the ith row take the value of 1 for observations neighboring the ith observation and 0 otherwise. Note that the matrix is symmetric and always has zeros on the lead diagonal. In applied work, a transformation is often used that converts the spatial matrix so that the rows sum to unity. This matrix is referred to as the standardized version of the spatial weight matrix.

Using the 'Delaunay triangle algorithm', contiguity information can be created artificially. This algorithm simply requires the latitude and longitude of each spatial unit. This procedure is also less time and computing intensive than using information on common borders and edges.[15] Given points in Cartesian space, the algorithm creates a set of triangles such that no points are contained in any triangle's circumcircle. The edges of the triangle satisfy the 'empty circle' property: the circumcircle of a triangle formed by three points is *empty* if if does not contain the vertices other than the three that define it. This way of creating a contiguity matrix entails more neighbors in relatively more densely populated areas and thus conforms to the idea that closer dwellings are more likely to be correlated with each other than those that are located at a distance from each other.

Once we have defined the contiguity or spatial weight matrix, spatial correlation between observations is captured in the error term u_{kth} in equation (7) as follows:

$$u_{kth} = \lambda W u_{kth} + \epsilon_{kth}, \tag{9}$$

where $\epsilon_{kth} \sim N(0, \omega_{kth}\sigma^2)$. The variance of ϵ_{kth} is subscripted with kt implying that the model will allow for heteroscedasticity. Furthermore, we assume that $\omega_{kth}\sigma^2 = g(x)$,

[14] A spatial matrix with binary numbers contains relatively less information but makes econometric estimation computationally less intensive.

[15] Matlab 6.5 has an in-built Delaunay triangle algorithm routine.

where x is a subset of the explanatory variables. W is a spatial weights matrix, and the parameter λ measures the average locational influence of the neighboring observations on each observations. For example, $\lambda = 0.30$ means that 30 per cent of the variation of u_{kth} is explained by locational influences of its neighbors.

3.3 Estimation of the model

The model is estimated separately for six sub-samples of the data set. Each of these samples contains observations for five consecutive quarters with the exception of the last sample where the observations belong to four quarters.[16] The existence of overlapping periods allows us to link the results from one block to the next.

Estimating equations separately for each five-quarter block allows the shadow prices of characteristics to vary over time. The shadow prices also vary in the spatial dimension because of the inclusion of interaction terms between regions and characteristics of dwellings. With this flexible set-up, we can construct temporal indexes for each region directly from the estimated coefficients, $\hat{\beta}_t$ and $\hat{\delta}_{kt}$. However, construction of spatial indexes is relatively more involved. This is because the interaction between regions and characteristics requires the inclusion of more coefficients in the calculation.

We use the maximum likelihood estimation (MLE) method developed by Anselin (1988) to estimate the parameters of the model.[17] We extend the method to account for heteroscedastic disturbance terms. The maximum likelihood method is based on a

[16] The six sub-samples cover the periods 2001:1–2002:1, 2002:1–2003:1, 2003:1–2004:1, 2004:1–2005:1, 2005:1–2006:1 and 2006:1–2006:4.

[17] See LeSage (1999) for a brief explanation of alternative spatial auto-regressive models. The most general model is $Y = \rho W_1 Y + X\beta + u$, where $u = \lambda W_2 u + \epsilon$ and $\epsilon \sim N(0, \sigma^2 I_n)$. W_1 and W_2 are two spatial contiguity matrices. Two of the alternative models can be specified by imposing restrictions on W_1 and W_2. By setting $W_1 = 0$, we specify a model which corrects for spatial correlation in the disturbance. The restriction of $W_2 = 0$ specifies a model that Anselin (1988) refers to as a mixed regressive-spatial autoregressive model. Our preliminary analysis shows that both models yield similar results. However, construction of indexes is computationally much less intensive in the former, which accounts for spatial correlation without yielding an additional slope coefficient.

concentrated likelihood function where $\hat{\lambda}$ is estimated from a univariate optimization routine. Other estimates of the model are obtained from ordinary least squares estimation. The process is conducted iteratively until the convergence criteria are satisfied. In each iteration, two estimations are undertaken, each updating its own estimates by using the updated estimates from the other estimation.

To start the process, we estimate an ordinary least squares (OLS) model: the log of dwelling price on the same set of variables specified in equation (7), with the assumption of homoscedastic error variance and $\lambda = 0$. The next step is to relax the assumption of homoscedasticity. We transform the model with the assumption that the error variance is a function of some of the explanatory variables, $g(x)$. Here x includes region, period, region-period interactions, postcodes and the core characteristics, but excludes all other interaction terms. The following transformation of the variables is taken: $Y^* = [1/\sqrt{g(x)}]Y$ and $X^* = [1/\sqrt{g(x)}]X$, where Y and X denote the dependent and explanatory variables of the model.

We run OLS on the transformed variables which gives us the feasible generalized least squares (FGLS) estimates of the intercept and slope coefficients. The FGLS estimates are fed into the following concentrated log likelihood function which is optimized over a single coefficient λ:

$$L(\lambda) = -\frac{n}{2}\ln(\pi) - \frac{n}{2}\ln(\frac{1}{n})\,[(I_n - \lambda W)\epsilon]'\,[(I - \lambda W)\epsilon] - \frac{n}{2} + \ln|I_n - \lambda W|. \qquad (10)$$

We again transform the variables but this time to correct for spatial correlation. The transformation takes place in the following way: $Y^{**} = [I_n - \tilde{\lambda}W]Y^*$ and $X^{**} = [I_n - \tilde{\lambda}W]X^*$, where I_n is an $n \times n$ identity matrix, $\tilde{\lambda}$ is the MLE from the current iteration and W is the spatial contiguity matrix defined earlier.

This completes one iteration. When the convergence criteria are satisfied, this iterative process yields a set of FGLS estimates for the intercept and the slope coefficients and maximum likelihood estimates for the spatial correlation coefficient. The FGLS coefficients, which we use later to construct indexes, are consistent and asymptotically efficient.[18]

4 Derivation of the Price Indexes

4.1 Temporal indexes

For a given region, k, temporal price indexes can be constructed from the estimated β and δ coefficients from the hedonic equation (7). Our price indexes are computed on a quarterly basis. Hence let $P_{k,t,u}$ denote the price index for region k in year t and quarter u. The price index $P_{k,t,u}$ relative to $P_{k,t,1}$ is calculated as follows:

$$\frac{P_{k,t,u}}{P_{k,t,1}} = \exp(\hat{\beta}_{t,u} + \hat{\delta}_{k,t,u}) \quad \text{for} \quad u = 2, 3, 4. \tag{11}$$

Similarly, the price index for the first quarter in year $t+1$ relative to the first quarter in year t is given by

$$\frac{P_{k,t+1,1}}{P_{k,t,1}} = \exp(\hat{\beta}_{t+1,1} + \hat{\delta}_{k,t+1,1}).$$

Comparisons between other pairs of quarters (say the second and third quarter) are obtained indirectly as follows:

$$\frac{P_{k,t,3}}{P_{k,t,2}} = \frac{P_{k,t,3}}{P_{k,t,1}} \times \frac{P_{k,t,1}}{P_{k,t,2}} = \exp[(\hat{\beta}_{t,3} - \hat{\beta}_{t,2}) + (\hat{\delta}_{k,t,3} - \hat{\delta}_{k,t,2})]. \tag{12}$$

Comparisons over longer time horizons require the linking of five-quarter blocks. For the case of chronologically adjacent blocks, the price indexes are calculated as follows:

$$\frac{P_{k,t+1,u}}{P_{k,t,1}} = \frac{P_{k,t+1,1}}{P_{k,t,1}} \times \frac{P_{k,t+1,u}}{P_{k,t+1,1}} = \exp[(\hat{\beta}_{t+1,1} + \hat{\beta}_{t+1,u}) + (\hat{\delta}_{k,t+1,1} + \hat{\delta}_{k,t+1,u})] \quad \text{for} \quad u = 2, 3, 4.$$

[18]There are a number of computational issues that are specific to spatial statistics/econometrics. The most important one is that it is necessary to compute the determinant of an $n \times n$ contiguity matrix, which requires a lot of computer memory. We have used the Spatial Econometrics Toolbox for MATLAB developed by J. P. LeSage available at ≪http://www.spatial-econometrics.com≫. The details on the toolbox and the associated computational issues can be found in LeSage (1999). For initial conditions, convergence criteria and maximum number of allowed iterations we have used the default settings of the Toolbox. Another toolbox for MATLAB which implements various spatial auto-regressive models has been developed by R. K. Pace and R. Barry ≪http://www.spatial-statistics.com≫.

For the more general case,

$$\frac{P_{k,t+s,1}}{P_{k,t,1}} = \frac{P_{k,t+1,1}}{P_{k,t,1}} \times \frac{P_{k,t+2,1}}{P_{k,t+1,1}} \times \cdots \times \frac{P_{k,t+s,1}}{P_{k,t+s-1,1}} = \prod_{j=1}^{s} \exp(\hat{\beta}_{t+j,1} + \hat{\delta}_{k,t+j,1}),$$

$$\frac{P_{k,t+s,u}}{P_{k,t,1}} = \exp(\hat{\beta}_{t+s,u} + \hat{\delta}_{k,t+s,u}) \prod_{j=1}^{s} \exp(\hat{\beta}_{t+j,1} + \hat{\delta}_{k,t+j,1}) \quad \text{for } \; u = 2, 3, 4. \tag{13}$$

Using this approach, it is possible to construct a temporal price index for each region over the entire period of analysis.

4.2 Spatial indexes

For a given quarter (t, u), spatial price indexes can be constructed from the estimated coefficients γ_k, δ_{kt}, η_{km} and θ_{kc} obtained from the hedonic equation (7).[19] Our starting point is a comparison between a postcode m in region j and a postcode n in region k for a particular dwelling h with characteristic vector z_{ch}. This spatial price index is calculated as follows:

$$P_{jmtu,kntu}(z_{ch}) = \exp[(\hat{\gamma}_k - \hat{\gamma}_j) + (\hat{\delta}_{kt} - \hat{\delta}_{jt}) + (\hat{\eta}_{kn} - \hat{\eta}_{jm})] \left[\prod_{c=1}^{C} \exp(\hat{\theta}_{kc} z_{ch} - \hat{\theta}_{jc} z_{ch}) \right] \tag{14}$$

The spatial index can be generalized to take account of all dwellings sold in postcodes jm as follows:

$$P^{L}_{jmtu,kntu} = \exp[(\hat{\gamma}_k - \hat{\gamma}_j) + (\hat{\delta}_{kt} - \hat{\delta}_{jt}) + (\hat{\eta}_{kn} - \hat{\eta}_{jm})] \left[\prod_{h=1}^{H_{jmtu}} \prod_{c=1}^{C} \exp(\hat{\theta}_{kc} z_{ch} - \hat{\theta}_{jc} z_{ch}) \right]^{1/H_{jmtu}} \tag{15}$$

The superscript L on the price index denotes the fact that it is analogous to a Laspeyres price index in the sense that they are calculated using the dwellings actually sold in postcode jm. In an analogous manner, Paasche-type indexes can be computed based on the dwellings actually sold in postcode kn as follows:

$$P^{P}_{jmtu,kntu} = \exp[(\hat{\gamma}_k - \hat{\gamma}_j) + (\hat{\delta}_{kt} - \hat{\delta}_{jt}) + (\hat{\eta}_{kn} - \hat{\eta}_{jm})] \left[\prod_{h=1}^{H_{kntu}} \prod_{c=1}^{C} \exp(\hat{\theta}_{kc} z_{ch} - \hat{\theta}_{jc} z_{ch}) \right]^{1/H_{kntu}} \tag{16}$$

[19]Spatial indexes are constructed for a given quarter. Therefore, β drops out of the calculation.

Taking the geometric mean of the Laspeyres and Paasche type indexes, we obtain a Fisher-type index that treats both postcodes symmetrically.

$$P^F_{jmtu,kntu} = (P^L_{jmtu,kntu} \times P^P_{jmtu,kntu})^{1/2} \tag{17}$$

A spatial index between pairs of regions (as opposed to postcodes) can now be obtained by comparing all possible combinations of postcodes between pairs of regions, and then taking the geometric mean of the results of all these pairwise comparisons, as follows:

$$P^F_{jtu,ktu} = \prod_{m=1}^{M_j} \prod_{n=1}^{M_k} (P^F_{jmtu,kntu})^{1/(M_j M_k)}, \tag{18}$$

where M_j and M_k denote the number of postcodes, respectively, in regions j and k.

These bilateral indexes are not transitive. That is, $P^F_{jtu,ktu} \times P^F_{ktu,ltu} \neq P^F_{jtu,ltu}$. These indexes can be transitivized using the Gini-EKS formula (see for example Hill 1997), as follows:

$$\frac{P_{ktu}}{P_{jtu}} = \prod_{i=1}^{K} \left(\frac{P^F_{ktu,itu}}{P^F_{jtu,itu}} \right)^{1/K}.$$

Transitive indexes are referred to as multilateral indexes in the price index literature (again see Hill 1997).

4.3 Panel indexes

Combining the temporal and spatial indexes of the previous sections to allow a price comparison between a region-period jsu and another region-period ktv, where neither the regions nor periods are matched, is not entirely straightforward. To see why, it is useful to represent the panel comparison as a graph in which each region-period is depicted by a vertex as shown in Figure 3 for the case of 9 quarters and 15 regions. Multilateral spatial benchmarks for 2001(1), 2002(1) and 2003(1), calculated using the Gini-EKS formula are represented by elongated ovals. If we try to combine more than one multilateral spatial benchmark with chained temporal indexes, each bilateral comparison within the panel will be path dependent. For example, a comparison between region-periods A-2001(1) and B-2002(4) in Figure 3 could be made by following a number of paths. Here we will consider just two. One path would entail linking a comparison between A-2001(1) and A-2002(4) with a comparison between A-2002(4) and B-2002(4). A second path would link a comparison between A-2001(1) and B-2001(1) with a comparison between B-2001(1) and B-2002(4). These two paths from A-2001(1) to B-2002(4) will yield different answers.

Insert Figure 3 Here

To obtain an internally consistent set of panel results it is necessary that all cycles are removed from the graph. This can be done in a number of ways (see Hill 2004). One simple way of doing this is to use only one spatial benchmark. Three such methods are depicted in Figure 4. An attractive feature of such methods is that they preserve the chained temporal indexes. The temporal indexes are generally considered more reliable than their spatial counterparts since they are probably less affected by omitted variables bias. A disadvantage of the methods in Figure 4 is that the paths between pairs of region-periods in the graph may become rather long. For example, suppose we use the spatial benchmark in 2001(1), and thereafter extrapolate forward using the chained temporal indexes. This would lead to some long paths between some pairs of vertices in the graph. For example, a comparison between region-periods A-2006(1) and B-2006(1) would follow the path: $A\text{-}2006(1) \rightarrow A\text{-}2005(1) \rightarrow \cdots \rightarrow A\text{-}2001(1) \rightarrow B\text{-}2001(1) \rightarrow B\text{-}2001(2) \rightarrow \cdots \rightarrow B\text{-}2006(1)$. The problem with this is that longer paths tend to cause the results to drift (due to the accumulation of errors). Also, the results depend heavily on a single spatial benchmark.

For these reasons, we prefer to generate multiple sets of panel results using in turn the spatial benchmarks for 2001(1), 2002(1) and 2003(1), as shown in Figure 4, and, by extension, panel results using also the benchmarks for 2004(1), 2005(1) and 2006(1).[20,21] This generates a total of six sets of panel results. We combine them

[20] As a consequence of our overlapping five-quarter blocks, we actually obtain two EKS spatial benchmarks for the first quarter of each year (except 2001). We combine these benchmarks by taking their geometric mean.

[21] We do not consider here the possibility of generating panel results from second, third or fourth quarter spatial benchmarks.

by taking a geometric mean to get our overall results. By construction, this method preserves all the chained temporal indexes. All the adjustments required for internal consistency are forced onto the less reliable spatial indexes. Also, it treats all regions symmetrically, and likewise all spatial benchmarks symmetrically. One weakness of this method, however, is that only the temporal indexes satisfy temporal fixity. That is, if a new year of data is added to the data set, the results for bilateral comparisons involving two regions will change.[22]

Insert Figure 4 Here

If temporal fixity is required, we recommend using the method described in Figure 5. This method combines multiple spatial benchmarks (at one year intervals) with chained temporal indexes. Cycles in the graph are prevented by omitting the temporal indexes for one quarter in each year for all except one link region. In Figure 5, region C acts as the link region. Four different graphs are obtained depending on which quarter each year is omitted (for all except the link region). The problem with each of these graphs is that the excluded quarter introduces a structural break in the temporal indexes. For example, in the top graph in Figure 5, the structural break will occur in comparisons between the fourth quarter and the first quarter of the following year. These breaks can be smoothed out by taking a geometric mean of the four sets of results. To ensure that all regions are treated symmetrically, 15 sets of results are generated using each region in turn as the link. The overall panel results treat all regions and all quarters symmetrically. The main problem with this method is that it distorts the temporal indexes. For this reason, except when temporal fixity is considered essential, we prefer the method describe in Figure 4.

Insert Figure 5 Here

[22]Some ways of imposing temporal fixity are discussed in Hill (2004).

5 An Application to Sydney, Australia (2001-6)

5.1 Regression results

The model for each period is estimated around 400 parameters. The estimated coefficients with information on whether they are significant or not are provided in the Appendix of Syed, Hill and Melser (2008). With only a few exceptions, the coefficients of the models have the expected signs. Around 88 percent of the coefficients are found to be significant at the 5 percent significance level. The coefficients are stable in the sense that there are not large differences in the value of the characteristics, or in the premium for living in a particular postcode, between periods. Although with all the interaction terms it is difficult to determine the impact of each of the characteristics on price, generally the impact is found to be in the expected direction. For example, units are found to be cheaper than houses, more bedrooms and bathrooms add to the value of a dwelling and the lot size is found to have a positive effect on the price (at a diminishing rate). In summary, the results are reasonably consistent with prior expectations and are robust to sample periods.

We have conducted likelihood ratio tests to check whether groups of variables are jointly significant. The log-likelihoods for each block of five quarters (or four quarters for the final block) are given in Table 4. Likelihood ratio (LR) test statistics are provided in Table 5. These statistics test one at a time a restriction against the most general model. For example, if we impose a restriction that the coefficients of all region-quarter dummies are zero, the LR test statistic is found to be 280.0 for the 2001(1)-2002(1) model. The LR test statistics have $\chi^2(k)$ distributions where k denotes the number of parameter restrictions. The tests show that the most important variables as a group are the postcodes. All the interaction terms are significant at the 1 percent level except for one case where it is significant at the 5 percent level. This implies that the use of the general model inclusive of interaction terms is justified.

Insert Table 4 Here

Insert Table 5 Here

Table 6 provides tests for spatial dependence and heteroscedasticity. As anticipated, the prices of dwellings exhibit a very high degree of spatial dependence. We regress each house price against the neighbors' house prices (no other variables), where neighbors are defined by a contiguity matrix constructed from a Delaunay triangulation algorithm. We find that around 70 percent of the house prices can be explained by the neighbors' prices. We have conducted some further tests to see how much of the spatial dependence is left out in the errors of the traditional hedonic model. Spatial dependence is reduced when locational variables such as postcodes and regions are included in the model. A further reduction is achieved when we account for spatial dependence explicitly in the model. However, around 20 percent of the spatial dependence still remains in the error. This analysis points us to two important facts: (1) hedonic regression models should take account of spatial dependence explicitly and (2) a simple spatial correction (as we did in our study) might not be enough to account for all the spatial dependence in the data.

Insert Table 6 Here

In order to check for the presence of heteroscedasticity, we conduct a Breusch-Pagan (BP) test on the residuals of the standard OLS model (see the notes below Table 6 for a description of the model and the test statistics). The BP statistics which have $\chi^2(k)$ distribution were significant for all the periods, which prompted us to estimate FGLS models.

The estimated model contains a lot of information including shadow prices of each of the characteristics, how they differ across regions, and how they interact with each other. Much of it is not our primary interest. Now we turn out attention to the resulting indexes obtained from the estimated coefficients.

5.2 Price indexes for a panel of 15 regions and 24 quarters

The panel indexes, calculated using the method described in Figure 4, are reported in Table 7 and their graphs are shown in Figure 6. These indexes can be used to make comparisons of prices across regions and time. The results are normalized such that region A (Inner Sydney) is equal to 1.00 in the first quarter of 2001. The interpretation of the numbers are that, for example, region N (Cronulla-Sutherland) in the first quarter of 2001 had prices which were only 71 percent of those in region A (Inner Sydney) in the same period, after controlling for physical and geo-spatial price determining characteristics. Turning first to the spatial dimension, there is clearly a great deal of disparity in the cost of housing across regions. A significant premium is being paid for dwellings in region F (Mosman-Cremorne) and region B (Eastern Suburb) and to a lesser extent in region C (Inner West). The results show that in the first quarter of 2001 the same dwelling in region D (Lower North Shore) would on average have cost a little more than double that in region K (Fairfield-Liverpool).

Insert Table 7 Here

Insert Figure 6 Here

Such premiums have two potential interpretations. First, they can be thought of as 'good' or 'bad' deals. People in region K are getting more for their money than those in expensive suburbs. This seems unlikely since these large price differentials persist over time. A second explanation is that the premiums embody unmeasured characteristics, reflecting everything that is left out of the hedonic function. That is, if we had access to all price determining characteristics then we would not expect to find such systematic premiums.

From Table 7 it can be seen that there are also quite significant differences in the price trends across regions. From 2001 to 2006, prices rose most (53.3 percent) in region O (Penrith-Windsor), and least (33.6 percent) in region F (Mosman-Cremorne).

5.3 A combined temporal index for the 15 regions

Our regional indexes can be combined to obtain an overall index for Sydney. We consider two ways of doing this. The first index is obtained by taking the geometric mean of the indexes for each region as follows:

$$\frac{P_{tv}}{P_{su}} = \prod_{k=1}^{K} \left(\frac{P_{ktv}}{P_{ksu}} \right)^{1/K},$$

where s and t denote years, while u and v denote quarters. The second index is calculated by taking a weighted average where weights correspond to the number of housing sales in the given period.[23]

$$\frac{P_{tv}}{P_{su}} = \prod_{k=1}^{K} \left(\frac{P_{ktv}}{P_{ksu}} \right)^{w_{k,su,tv}},$$

where

$$w_{k,su,tv} = \frac{H_{ksu} + H_{ktv}}{\sum_{j=1}^{K} (H_{jsu} + H_{jtv})}.$$

Both indexes depicted in Table 8 indicate that prices increased between the first quarter of 2001 and the last quarter of 2003, after which they declined until the end of the sample period (December 2006). This finding is consistent with the consensus view (see Robertson 2006).

Insert Table 8 Here

In the period of increasing prices, both the hedonic indexes reported in Table 8 are for the most part indistinguishable, exhibiting the same rate of price increase – about 50 percent in three years. When prices started decreasing, the first index exhibits a faster rate of decrease than the weighted mean index. This indicates that the regions with a higher volume of sales had a slower decline than the regions which had lower sales.

Also included in Table 8, is the Australian Bureau of Statistics (ABS) house price index for Sydney.[24] The three indexes are graphed in Figure 7. Before considering the difference in trends, we should point out some similarities. All the series indicate the same 'boom' period for the housing market, from the first quarter of 2001 to fourth quarter of 2003. The market has stabilized and started falling since then, which is

[23] We have also constructed an index by taking value weights and found it almost indistinguishable from the index obtained using 'number of sales' weights.

[24] The series in ABS (2003) had a different base than the series in ABS (2007). We have changed the base to link up the two series so as to allow a comparison with our series.

again indicated by both the ABS and our hedonic indexes. Prices increased as much as 50 percent in some regions between 2001 and 2003. Since then they have fallen only slightly.

Insert Figure 7 Here

While the broad trends in our hedonic indexes and the ABS index are similar, there are also important differences between the two sets of results. In particular, the ABS index rose by more than 60 percent from 2001 to 2003, while our hedonic indexes rose by only about 50 percent. The ABS indexes also fell more from 2003 onwards. In fact, the fall in house prices after 2003 is only just discernable in our hedonic indexes in Figure 7. Also, the ABS index seems to be more volatile. These differences can provide interesting insights. The ABS index is a stratified median price index. The sample is stratified based on location. Each strata (or cluster) consisting of houses of similar price determining characteristics is formed using principal component analysis. The index for a city is obtained by taking a weighted average of the 'median price' of each strata. The ABS series includes only 'project homes' and 'established houses', essentially residential dwellings on their own block of land.[25] The ABS index excludes apartment housing and high-rise buildings, and hence covers a narrower range of dwellings than our indexes. The difference may be suggestive of the fact that established houses, as defined by the ABS, behaved differently from the rest of the market. Alternative explanations include differences in the source and scope of data, sampling methods and methodological differences in compiling and calculating the indexes.

With regard to methodological differences, even if they start from identical data sets, average price methods, such as the ABS stratified median method, and hedonic methods may generate quite different results. If properly modelled, a hedonic method reflects a 'pure price' change, abstracting from qualitative and compositional changes, whereas an 'average price' method combines the two changes. In boom periods, the premium from each additional attribute is expected to be higher. Therefore, it is likely

[25]See ABS (2005) for details on the ABS House Price Index and on some recent efforts to revamp the series.

that new dwellings that entered the market were increasingly of better quality (larger, with more bedrooms and bathrooms) and old dwellings were renovated. This might also explain the divergence between the ABS and our hedonic indexes. In the post-boom period, the ABS index fell faster than our hedonic indexes. This is suggestive of the fact that lower quality and/or smaller dwellings may have come on the market earlier.

5.4 Spatial indexes and convergence

Spatial indexes might be of interest for several reasons. For a given period, they provide a ranking of the regions in terms of the value of houses. This has policy implications because it indicates which civic amenities (such as parks, footpaths, shopping areas) are valued most, providing guidance for future public expenditure, as well as providing some indication of the degree of inequality of wealth. Spatial indexes can also shed light on the extent of market segmentation and regional variations in the impact of business cycles. This latter issue is particularly relevant here given that house prices peaked in 2003.

Spatial indexes obtained from equation (18) are provided in Table 9. From Table 9 it is clear that there are significant and systematic differences in prices across regions. Prices in the most expensive region (Region B: Eastern Suburbs) are more than three times higher than in the least expensive region (Region O: Penrith-Windsor) and, though it narrowed slightly, the gap remained throughout the whole period of analysis. Also, with few exceptions, the ranking of the regions remains the same over time.

Insert Table 9 Here

To investigate whether differences in price levels across regions are rising or falling over time, we calculate σ-convergence coefficients for the 15 regions in each of the 24 quarters in our data set. σ-convergence measures the variance of the cross-section of price parities and then examines whether this has declined or increased over time (see for example Sala-i-Martin 1996). That is, we calculate and compare the following:

$$\sigma_t^2 = \frac{1}{K}\sum_{k=1}^{K}\left[\ln(P_{kt}) - \ln(\bar{P}_t)\right]^2, \quad \ln(\bar{P}_t) = \frac{1}{K}\sum_{k=1}^{K}\ln(P_{kt}), \quad t = 1,\ldots,T. \qquad (19)$$

Applying this formula to the multilateral indexes in Table 7, we find a clear pattern of convergence until early 2004 (when house prices were rising) followed by divergence thereafter (when house prices were falling). The sigma convergence results are shown in Table 10. The convergence turning point appears to lag the change in direction of the housing market by one or two quarters. This association of the house price movement and the measure of σ-convergence is illustrated graphically in Figure 8.

Insert Table 10 Here

Insert Figure 8 Here

One possible explanation for this finding is that a rise in house prices in the richest regions triggered by a scarcity of housing in desirable locations (such as close to Sydney Harbor) creating a perception that house prices were rising throughout Sydney, and that this perception then became self-fulfilling. That is, price rises in poorer regions were driven more by momentum than fundamentals, as compared with richer regions. When the market started to decline, it follows that prices in the poorer regions fell more. It remains to be seen whether similar patterns are observed in other cities.

6 Conclusion

In this study we have constructed panel price indexes for 15 regions in Sydney, Australia over 24 quarters using a hedonic regression model. Our hedonic model is flexible in that it includes interaction terms, and allows the shadow prices on characteristics to evolve over time. The latter is achieved by breaking the comparison up into blocks and then using innovative methods to splice the blocks together. We also account for missing characteristics, spatial correlation and heteroscedasticity, all of which are prevalent in housing data sets. In particular, ours is the first study to apply the multiple-imputation method to the missing-data problem in a housing context.

The model performed well in terms of the economic and statistical significance of the parameters. The likelihood ratio tests on groups of variables show that the interaction between characteristics and regions is significant indicating that the inclusion of interaction terms is justified.

Our hedonic house price indexes rose significantly from 2001 to 2003, after which they fell slightly. This finding is consistent with the official Australian Bureau of Statistics (ABS) index. Our indexes, however, are less volatile than their ABS counterpart, rising noticeably less in the boom and falling less thereafter. We discuss some possible causes of these differences. In the spatial dimension, we find large and systematic differences in the price of housing across regions. The regional dispersion narrowed during the boom period but appears to have increased again since then. It remains to be seen whether a similar pattern will be observed in other cities.

House price indexes are important since many consumers and investors, and also government, in some way or another are tied to the housing market. They are also an important input into the overall measure of inflation. In addition to providing a better measure of 'pure price' change over time in the housing market, our study also considers regional variations. Our findings may have policy implications for macroeconomic management and resource allocation at the regional level. Finally, although the focus was on the housing market, our methodology may also be usefully applied in other markets requiring qualitative and compositional adjustments.

References

Aizcorbe A. and B. Aten (2004),"An Approach to Pooled Time and Space Comparisons," Paper presented at the SSHRC Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver.

Allison P. D. (2002), *Missing Data*, Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-136, Thousand Oaks, CA: Sage.

Anselin L. (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwar Academic Publishers.

Australian Bureau of Statistics (2007), *House Price Indexes: Eight Capital Cities*, September Quarter 2007 (released 7 November 2007), Catalogue no. 6416.0.

Australian Bureau of Statistics (2005), *Information Paper: Renovating the Established House Price Index*, (released November 2005), Catalogue No. 6417.0.

Australian Bureau of Statistics (2003), *House Price Indexes: Eight Capital Cities*, December Quarter 2003 (released 4 March 2004), Catalogue no. 6416.0.

Basu S. and T. Thibodeau (1998), "Analysis of Spatial Autocorrelation in House Prices," *Journal of Real Estate Finance and Economics* 17(1), 61-85.

Berndt E. R., Z. Griliches and N. J. Rappaport (1995), "Econometric Estimates of Price Indexes for Personal Computers in the 1990's," *Journal of Econometrics* 68(1), 243-268.

Bourassa S. C., M. Hoesli and V. C. Peng (2003), "Do Housing Submarkets Really Matter?" *Journal of Housing Economics* 12(1), 12-18.

Can A. (1990), "The Measurement of Neighbourhood Dynamics in Urban Housing Prices," *Economic Geography* 66(3), 254-272.

Case K. E., J. M. Quigley and R. J. Shiller (2005), "Comparing Wealth Effects: The Stock Market versus the Housing Market," *Advances in Macroeconomics* 5(1), 1-32.

Collins L. M., J. L. Schafer and C. M. Kam (2001), "A comparison of inclusive and restrictive strategies in modern missing-data procedures," *Psychological Methods* 6, 330-351.

Court A. T. (1939), "Hedonic Price Indexes with Automotive Examples," in *The Dynamics of Automobile Demand*. The General Motors Corporation, New York, 99-117.

Diewert W. E., S. Heravi and M. Silver (2007), "Hedonic Imputation versus Time Dummy Hedonic Indexes," IMF Working Paper No. 07/234, Washington D.C.

Englund P., J. M. Quigley and C. L. Redfearn (1998),"Improved Price Indexes for Real Estate: Measuring the Course of Swedish House Prices," *Journal of Urban Economics* 44, 181-196.

Garderen K. J. van and C. Shah (2002), “Exact Interpretation of Dummy Variables in Semilogarithmic Equations,” *Econometrics Journal* 5, 149-159.

Griliches Z. (1961), “Hedonic Price Indexes for Automobiles: An Econometric Analysis for Quality Change,” In Z. Griliches (eds.), *Price Indexes and Quality Change: Studies in New Methods of Measurement*, Cambridge: Harvard University Press.

Hansen J. (2006),“Australian House Prices: A Comparison of Hedonic and Repeat Sales Measures,” Reserve Bank of Australia Discussion Paper, No. 2006-03.

Hill R. J. (2004),“Comparing Price Indexes Across Countries and Time: The Case of the European Union,” *American Economic Review* 94(5), 1379-1410.

Hill R. J. (1997), “A Taxonomy of Multilateral Methods for Making International Comparisons of Prices and Quantities,” *Review of Income and Wealth* 43(1), 49-69.

Hill R. J. and D. Melser (2008a),“Hedonic Imputation the Price Index Problem: An Application to Housing,” *Economic Inquiry*, forthcoming.

Hill R. J. and D. Melser (2008b),“Comparing House Prices Using Hedonic Methods: The Problem of Substitution Bias,” Mimeo.

International Labour Office (ILO)(2004), *Consumer Price Index Manual: Theory and Practice*, Geneva: ILO Publications.

Kelejian H. H. and D. P. Robinson (1995), “Spatial Correlation: A Suggestive Alternative to the Autoregressive Model,” In L. Anselin and R. J. G. M. Florax (eds.), *New Directions in Spatial Econometrics*. Berlin: Springer.

Kennedy P. E. (1981),“Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations,” *American Economic Review* 71(4), 801.

Lancaster K. J. (1966),“A New Approach to Consumer Theory,” *The Journal of Political Economy* 74 (2), 132-157.

LeSage J. P. (1999), *The Theory and Practice of Spatial Econometrics*, Department of Economics, University of Toledo, Ohio.

Little R. J. A. and D. B. Rubin (1987), *Statistical Analysis with Missing Data*, New York: Wiley.

Pace R. K., R. Barry, J. M. Clapp and M. Rodriguez (1998), "Spatiotemporal autoregressive models of neighborhood effects," *Journal of Real Estate Finance and Economics* 17(1), 15-33.

Pace R. K., R. Barry and C. F. Sirmans (1998), "Spatial Econometrics and Real Estate," *Journal of Real Estate Finance and Economics* 17, 5-13.

Pace R. K. and O. Gilley (1997), "Using Spatial Configuration of the Data to Improve Estimation," *Journal of Real Estate Finance and Economics* 14(3), 333-340.

Pakes A. (2003), "A Reconsideration of Hedonic Price Indexes with an Application to PCs," *American Economic Review* 93, 1576-1593.

Robertson R. (2006), *Thinking About the Big Drop in Australian Housing Affordability*, Discussion Draft, Macquarie Research Strategy, Macquarie Bank.

Rosen S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82(1), 34-55.

Rubin D. B. (1996), "Multiple Imputation After 18+ Years," *Journal of the American Statistical Association* 91, 473-489.

Rubin D. B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley.

Rubin D. B. (1976), "Inference and Missing Data," *Biometrika* 63, 581-592.

Sala-i-Martin X (1996), "The Classical Approach to Convergence Analysis," *The Economic Journal* 106 (437), 1019-1036.

Schafer J. L. (1999), "Multiple imputation: A primer," *Statistical Methods in Medical Research*, 8, 3-15.

Schafer J. L. (1997), *Analysis of Incomplete Multivariate Data*, London: Chapman & Hall.

Schafer J. L. and J. W. Graham (2002), "Missing Data: Our View of the State of the Art," *Psychological Methods* 7(2), 147-177.

Triplett J. E. (2004), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*, STI Working Paper 2004/9, Directorate for Science, Technology and Industry, Organisation for Economic Co-operation and Development, Paris.

Yuan Y. C. (2000), "Multiple Imputation for Missing Data: Concepts and New Development," *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference* Cary, NC: SAS Institute Inc., 1410-1419.

Table 1: Some Summary Information of Dwelling Prices

	Number of Observations	Price (Australian $) Mean	Median	Std Dev.
Full sample	418,877	527,095	422,000	381,045
By type of dwelling:				
Houses	222,033	602,863	495,000	413,377
Units	196,844	441,631	370,000	320,069
By bedroom counts:				
1	18,551	341,416	318,000	183,830
2	80,616	449,471	400,000	253,006
3	94,628	594,382	500,000	370,982
4	42,853	775,097	638,000	490,237
5 & above	11,735	988,669	800,000	627,499
By year:				
2001	80,233	431,064	350,000	322,896
2002	85,881	493,967	400,000	352,098
2003	80,741	551,110	445,000	387,192
2004	59,194	574,326	462,500	398,391
2005	58,187	572,320	460,000	403,375
2006	54,641	585,360	465,000	417,474
By region:				
A	33,484	537,063	435,000	393,578
B	34,432	785,706	595,000	566,461
C	24,788	598,312	525,000	340,999
D	23,973	717,292	580,000	466,254
E	22,291	707,260	625,000	399,452
F	9,561	838,470	592,000	636,903
G	24,354	694,801	595,000	442,679
H	24,248	559,706	518,000	300,293
I	35,818	452,371	377,000	308,421
J	32,033	407,387	369,000	330,000
K	39,110	336,580	320,000	190,154
L	17,555	348,669	325,000	185,112
M	34,209	443,500	403,000	215,871
N	25,715	548,257	480,000	323,908
O	37,306	312,076	299,000	133,583

Notes: (1) Because of missing data, the number of observations by bedroom counts does not add up to the total sample.

(2) A=Inner Sydney, B=Eastern Suburbs, C=Inner West, D=Lower North Shore, E=Upper North Shore, F=Mosman-Cremorne, G=Manly-Warringah, H=North Western, I=Western Suburbs, J=Parramatta Hills, K=Fairfield-Liverpool, L=Canterbury-Bankstown, M=St George, N=Cronulla-Sutherland & O=Penrith-Windsor.

Table 2: Percentage of Missing Data

	% of Observations Missing the Attribute(s):			
	Bed	Bath	Bed & Bath	Bed or Bath or lot size
Full sample	40.7	58.4	40.5	58.8
By type of dwelling:				
Houses	33.3	53.1	33.3	53.6
Units	49.0	64.3	48.7	64.6
By bedroom counts:				
1	n.a	30.4	n.a	30.5
2	n.a	34.0	n.a	34.3
3	n.a	31.1	n.a	31.6
4	n.a	22.7	n.a	23.2
5 & above	n.a	21.6	n.a	22.2
By year:				
2001	46.6	69.3	46.4	69.6
2002	48.8	68.5	48.7	68.8
2003	45.6	65.4	45.4	65.7
2004	39.5	61.6	39.3	62.0
2005	31.6	41.6	31.4	42.1
2006	23.1	30.4	22.8	31.3
By region:				
A	56.6	37.3	55.5	68.6
B	23.3	40.3	23.0	41.0
C	26.2	44.8	25.9	45.5
D	33.3	48.4	33.1	49.0
E	25.7	43.5	25.7	43.8
F	23.6	40.4	23.4	40.8
G	29.4	49.5	29.3	49.8
H	28.5	50.8	28.4	51.0
I	43.2	62.3	43.2	62.6
J	51.2	62.9	51.2	63.1
K	59.0	73.2	59.0	73.4
L	49.9	71.6	49.9	71.9
M	34.7	64.6	34.6	64.9
N	35.9	62.9	35.9	63.4
O	59.3	68.2	59.2	68.4

Note: In the case of many units, the available lot size information was for the whole floor the unit belonged to rather than only for that particular unit. Therefore lot size data was not used for units. However, they were not considered as missing. For houses, only 2150 observations had missing lot size.

Table 3: Percentage of Missing Data in the Most and Least Expensive Postcodes

Postcodes	Median Price ('000 $)	Observations	% of Observations Missing the Attribute(s): Bed	Bath	Bed & Bath	Bed or Bath or lot size
Least Expensive 5:						
2195	205	3254	47.8	75.2	47.7	75.4
2770	231	5656	50.9	59.2	50.8	59.4
2166	242	5254	71.2	86.5	71.2	86.7
2163	248	652	60.0	79.6	59.8	79.8
2760	251	3305	45.6	61.2	45.5	61.5
Most Expensive 5:						
2071	935	1233	11.9	25.9	11.8	26.0
2030	950	1486	19.7	35.3	19.6	35.8
2069	950	1282	13.9	30.0	13.9	30.3
2063	1050	756	19.3	33.9	19.0	34.1
2108	1170	343	23.0	37.9	23.0	38.2

Table 4: Log Likelihoods of the Estimated Equations

	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Observations	96837	100913	91067	70285	70756	54172
Parameters	404	404	404	404	404	389
Log Likelihood	443×10^3	470×10^3	419×10^3	314×10^3	313×10^3	232×10^3

Table 5: Likelihood Ratio Tests

Restrictions in Variables	No.Param Restrict	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Postcodes	175	10155.0	8478.4	7838.2	5815.9	6737.8	6631.0
Region-Quarter	56	280.0	401.4	253.8	185.3	155.7	124.9
All interactions with characteristics	144	6148.0	7028.0	7355.5	5319.8	5902.0	3787.2
Region-Dwell type	14	279.0	316.1	429.3	368.6	338.5	197.5
Region-Bedroom	56	587.0	585.3	540.6	472.9	530.5	442.8
Region-Bathroom	42	643.0	573.1	606.5	539.5	475.4	304.3
Region-Lot size	14	234.0	147.0	158.7	74.9	125.3	138.9
Dwelling type-Bed	4	85.0	207.3	157.2	132.6	303.8	107.5
Dwelling type-Bath	3	207.0	381.5	400.5	159.7	183.6	100.6
Bedroom-Lot size	4	46.0	22.5	27.2	39.4	42.5	15.3
Bathroom-Lot size	3	71.0	91.4	72.6	36.3	39.7	15.2
Bedroom-Bathroom	4	12.0	28.1	43.2	42.3	33.2	50.2

Notes: (1) The log-likelihood ratios test against the model specified in equation 7 have $\chi^2(k)$ distribution where k is the number of parameter restrictions imposed on the model. All the model restrictions are found to be binding at the 1% significance level except for the bedroom-bathroom interaction terms in period 1 which are jointly significant at the 5% level.

(2) The parameter restrictions for the region-quarter interaction terms are 42 for period 6.

Table 6: Spatial Correlation and Heteroscedasticity Tests

	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Spatial Correlation Tests:						
On Observed Dwelling Price	0.73	0.68	0.67	0.66	0.68	0.69
On Residuals from Models:						
Model a	0.43	0.40	0.38	0.38	0.40	0.39
Model b	0.34	0.33	0.32	0.30	0.31	0.29
Model c	0.11	0.22	0.20	0.19	0.19	0.18
Spatial Correlation Coefficient in Model c ($\hat{\lambda}$)	0.29	0.27	0.26	0.25	0.26	0.23
Breusch-Pagan(BP) F Stat.:	41	39	56	50	45	52

Notes: (1) Model a: A traditional hedonic model without consideration of spatial heterogeneity or the possibilities for the existence of sub-markets across spatial dimensions. This model excludes the postcode and region dummies, and any interactions with them, from equation (7). Additionally, it does not account for spatial correlation.

(2) Model b: It includes the postcode and region dummies and interactions with them. It does not account for spatial correlation.

(3) Model c: The model specified in equation (7).

(4) All the BP F Statistics are significant at the 1% level.

Table 7: Panel House Price Indexes for Sydney by Region

Quarter/ Region	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
2001 Q1	1.00	1.36	1.23	0.98	0.68	1.31	0.82	0.61	0.74	0.56	0.47	0.54	0.65	0.71	0.45
Q2	1.09	1.44	1.30	1.06	0.71	1.39	0.87	0.67	0.80	0.59	0.50	0.57	0.72	0.72	0.46
Q3	1.15	1.54	1.39	1.08	0.75	1.40	0.93	0.70	0.84	0.61	0.52	0.61	0.73	0.74	0.49
Q4	1.16	1.53	1.39	1.11	0.78	1.48	0.99	0.72	0.89	0.64	0.56	0.64	0.77	0.78	0.52
2002 Q1	1.19	1.58	1.38	1.12	0.80	1.55	1.01	0.75	0.94	0.67	0.59	0.68	0.79	0.81	0.55
Q2	1.27	1.68	1.50	1.20	0.83	1.53	1.04	0.81	1.00	0.70	0.64	0.71	0.84	0.86	0.58
Q3	1.25	1.75	1.51	1.22	0.87	1.58	1.05	0.83	1.04	0.75	0.68	0.76	0.89	0.87	0.61
Q4	1.28	1.72	1.59	1.24	0.89	1.54	1.09	0.85	1.07	0.78	0.72	0.78	0.89	0.92	0.64
2003 Q1	1.26	1.81	1.55	1.28	0.89	1.58	1.11	0.84	1.12	0.79	0.75	0.78	0.91	0.95	0.65
Q2	1.32	1.84	1.66	1.28	0.92	1.68	1.13	0.90	1.13	0.81	0.76	0.81	0.93	0.97	0.68
Q3	1.33	1.92	1.75	1.33	0.98	1.71	1.17	0.93	1.14	0.86	0.80	0.87	1.00	1.03	0.72
Q4	1.37	1.90	1.70	1.31	0.99	1.70	1.21	0.95	1.19	0.89	0.82	0.89	1.01	1.05	0.75
2004 Q1	1.34	1.87	1.70	1.28	0.98	1.70	1.18	0.94	1.15	0.88	0.83	0.86	1.00	1.06	0.76
Q2	1.36	1.80	1.67	1.24	0.96	1.62	1.17	0.92	1.12	0.86	0.82	0.87	0.98	1.02	0.75
Q3	1.37	1.86	1.68	1.30	0.98	1.64	1.17	0.91	1.14	0.85	0.82	0.87	0.97	0.99	0.75
Q4	1.35	1.87	1.66	1.37	0.98	1.68	1.20	0.93	1.15	0.84	0.82	0.85	0.96	1.01	0.76
2005 Q1	1.44	1.88	1.65	1.35	0.97	1.75	1.18	0.90	1.12	0.83	0.80	0.83	0.95	1.00	0.75
Q2	1.37	1.81	1.68	1.31	0.96	1.75	1.18	0.91	1.16	0.86	0.81	0.88	0.94	0.99	0.73
Q3	1.34	1.84	1.64	1.30	0.96	1.71	1.18	0.90	1.09	0.81	0.77	0.83	0.93	0.97	0.72
Q4	1.36	1.83	1.70	1.29	0.98	1.77	1.20	0.88	1.10	0.81	0.76	0.83	0.93	0.96	0.72
2006 Q1	1.38	1.81	1.68	1.29	0.97	1.74	1.18	0.90	1.10	0.81	0.75	0.80	0.92	0.97	0.72
Q2	1.38	1.86	1.71	1.33	0.98	1.82	1.17	0.91	1.08	0.81	0.75	0.80	0.92	0.97	0.71
Q3	1.33	1.88	1.68	1.33	0.97	1.77	1.22	0.89	1.08	0.80	0.72	0.77	0.93	0.95	0.72
Q4	1.38	1.91	1.71	1.31	0.99	1.75	1.25	0.90	1.05	0.80	0.70	0.76	0.91	0.96	0.69
Total Change(%)	38.0	40.4	39.0	33.7	45.6	33.6	52.4	47.5	41.9	42.9	48.9	40.7	40.0	35.2	53.3

Table 8: Temporal Price Indexes for Sydney

Quarters	Calculated from Panel Indexes Reported in Table 7		ABS
	Geometric Mean	Weighted Mean	Index
2001 Q1	1.00	1.00	1.00
Q2	1.06	1.06	1.03
Q3	1.11	1.12	1.09
Q4	1.16	1.16	1.15
2002 Q1	1.20	1.20	1.20
Q2	1.26	1.28	1.29
Q3	1.31	1.32	1.35
Q4	1.34	1.36	1.40
2003 Q1	1.37	1.36	1.42
Q2	1.41	1.41	1.48
Q3	1.47	1.47	1.56
Q4	1.50	1.50	1.62
2004 Q1	1.48	1.50	1.61
Q2	1.45	1.47	1.55
Q3	1.46	1.47	1.54
Q4	1.47	1.48	1.55
2005 Q1	1.46	1.49	1.51
Q2	1.46	1.49	1.49
Q3	1.43	1.45	1.47
Q4	1.43	1.46	1.48
2006 Q1	1.43	1.46	1.47
Q2	1.44	1.46	1.49
Q3	1.42	1.45	1.50
Q4	1.42	1.45	1.50
Total Change(%)	42.1	45.1	50.2

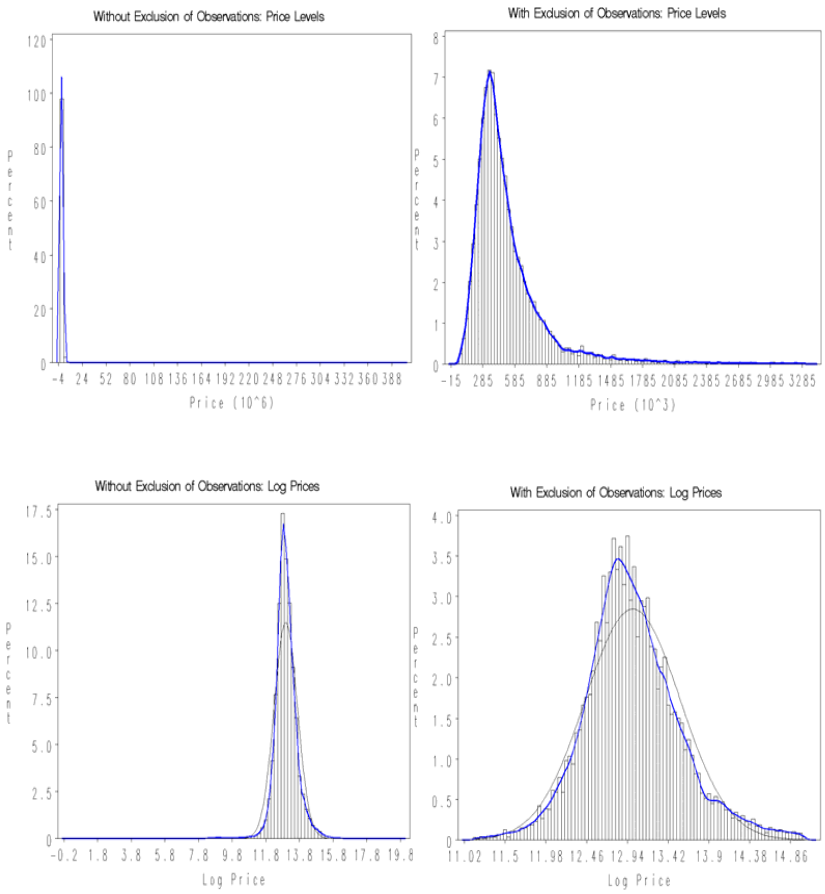
Table 9: Spatial Price Indexes for Sydney

Quarter/ Region	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
2001 Q1	1.00	1.36	1.23	0.98	0.68	1.31	0.82	0.61	0.74	0.56	0.47	0.54	0.65	0.71	0.45
2002 Q1	1.00	1.33	1.17	0.95	0.68	1.31	0.86	0.64	0.79	0.56	0.50	0.57	0.67	0.69	0.46
2003 Q1	1.00	1.43	1.22	1.01	0.70	1.25	0.87	0.67	0.89	0.62	0.59	0.62	0.72	0.75	0.52
2004 Q1	1.00	1.39	1.27	0.96	0.73	1.27	0.88	0.70	0.86	0.66	0.62	0.64	0.75	0.79	0.57
2005 Q1	1.00	1.31	1.14	0.93	0.67	1.22	0.82	0.62	0.78	0.58	0.56	0.58	0.66	0.70	0.52
2006 Q1	1.00	1.31	1.22	0.94	0.70	1.26	0.86	0.65	0.79	0.59	0.54	0.58	0.67	0.70	0.52

Table 10: Estimates of σ Convergence

Quarter		Variance	Quarter		Variance
2001	Q1	0.1208	2004	Q1	0.0728
	Q2	0.1248		Q2	0.0695
	Q3	0.1228		Q3	0.0748
	Q4	0.1143		Q4	0.0773
2002	Q1	0.1065	2005	Q1	0.0851
	Q2	0.1030		Q2	0.0788
	Q3	0.0951		Q3	0.0846
	Q4	0.0867		Q4	0.0896
2003	Q1	0.0870	2006	Q1	0.0900
	Q2	0.0892		Q2	0.0967
	Q3	0.0833		Q3	0.0981
	Q4	0.0746		Q4	0.1044

Figure 1: Distribution of Dwelling Prices Before and After Exclusion of Extreme Observations



Notes: (1) Observations excluded are: (i) top and bottom 1 per cent of the distribution of prices, (ii) top and bottom 1 per cent of the distribution of *lot size*, (iii) dwellings with more than 10 bedrooms and (iv) dwellings with more than 8 bathrooms.
(2) The smooth curves drawn on the histogram for price levels are the normal kernel density functions estimated from the data. Kernel density estimation is a non-parametric technique that averages a kernel function across observations with pre-specified bandwidth to create a smooth approximation of the distribution.
(3) In each of the histograms for the logarithm of prices, there is an additional smooth curve (lighter of the two in each diagram). These are the normal curves with means and standard deviations obtained from the empirical distributions.

Figure 2: Spatial Dependence of Prices of Dwellings

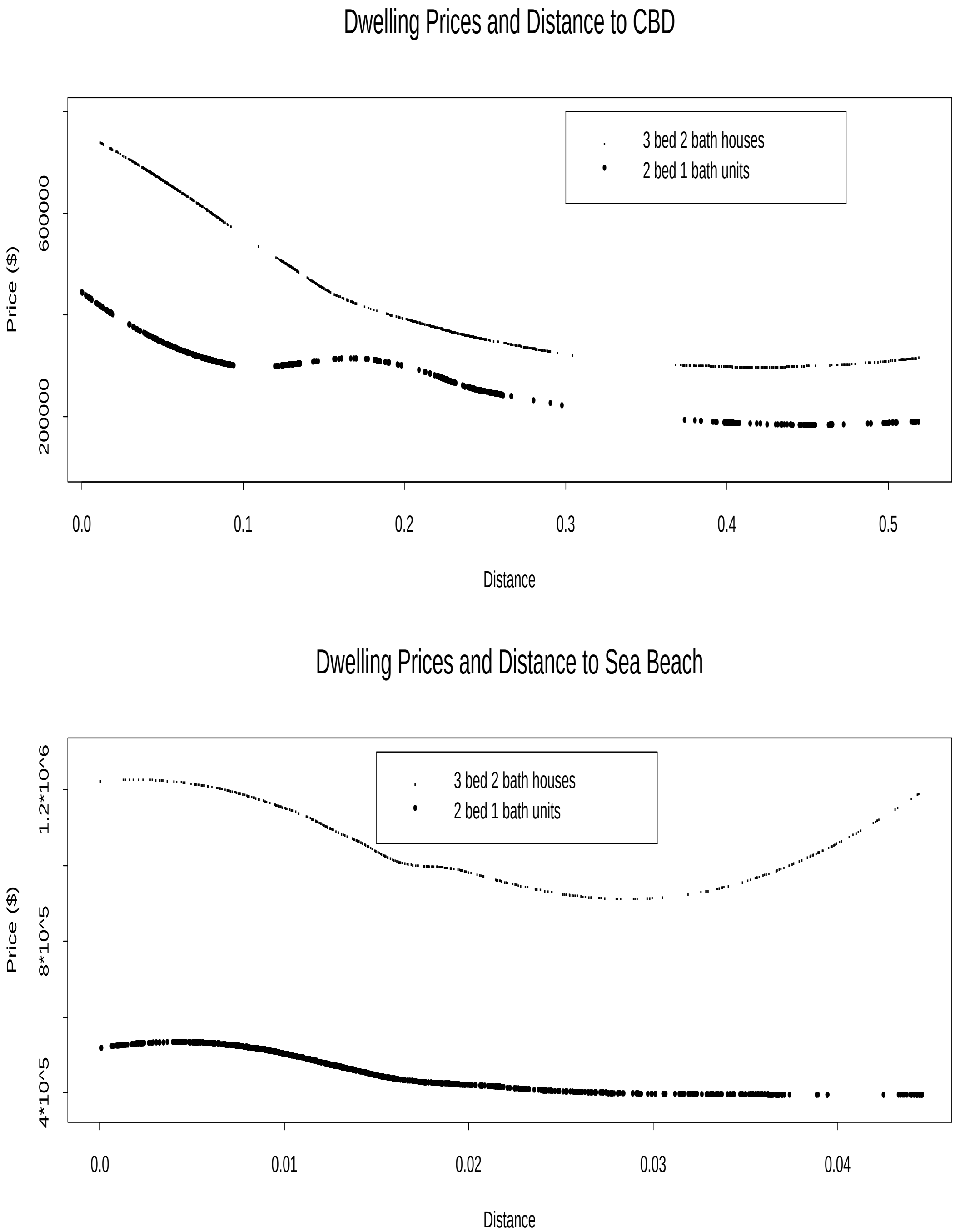
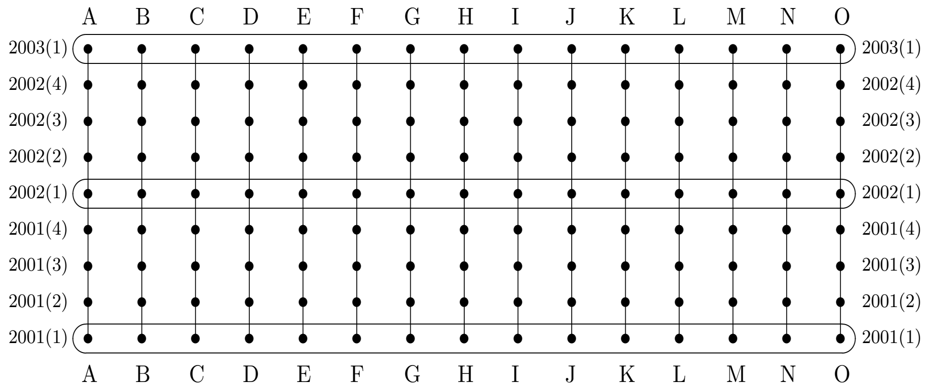


Figure 3: The Path Dependence Problem in Panel Comparisons



Note: The Residex regions used are: A=Inner Sydney, B=Eastern Suburbs, C=Inner West, D=Lower North Shore, E=Upper North Shore, F=Mosman-Cremorne, G=Manly-Warringah, H=North Western, I=Western Suburbs, J=Parramatta Hills, K=Fairfield-Liverpool, L=Canterbury-Bankstown, M=St George, N=Cronulla-Sutherland, Campbelltown, O=Penrith-Windsor.

Figure 4: Panel Comparisons That Use A Single Spatial Benchmark

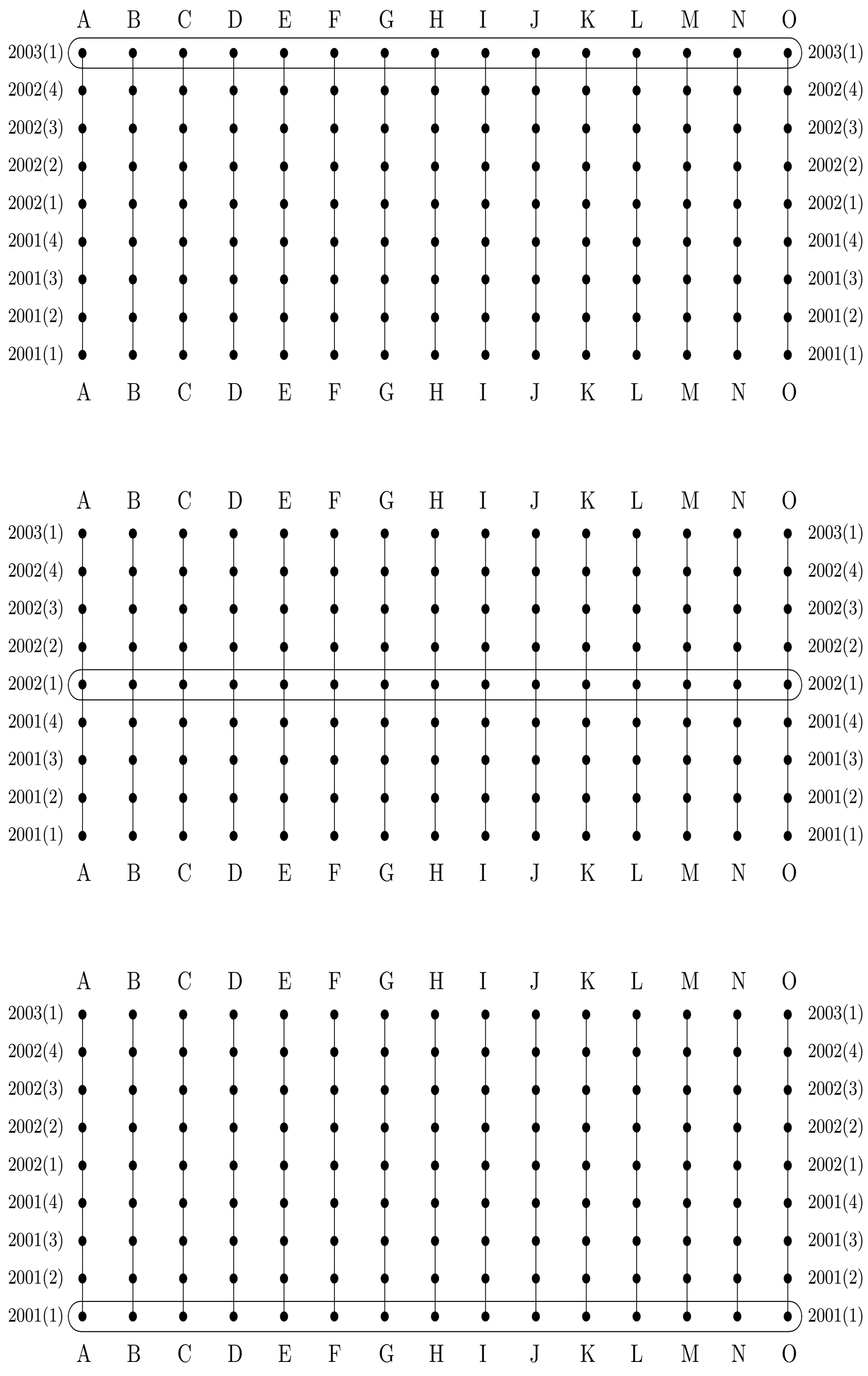


Figure 5: Panel Comparisons That Use Spatial Benchmarks at One-Year Intervals Linked Through Region C

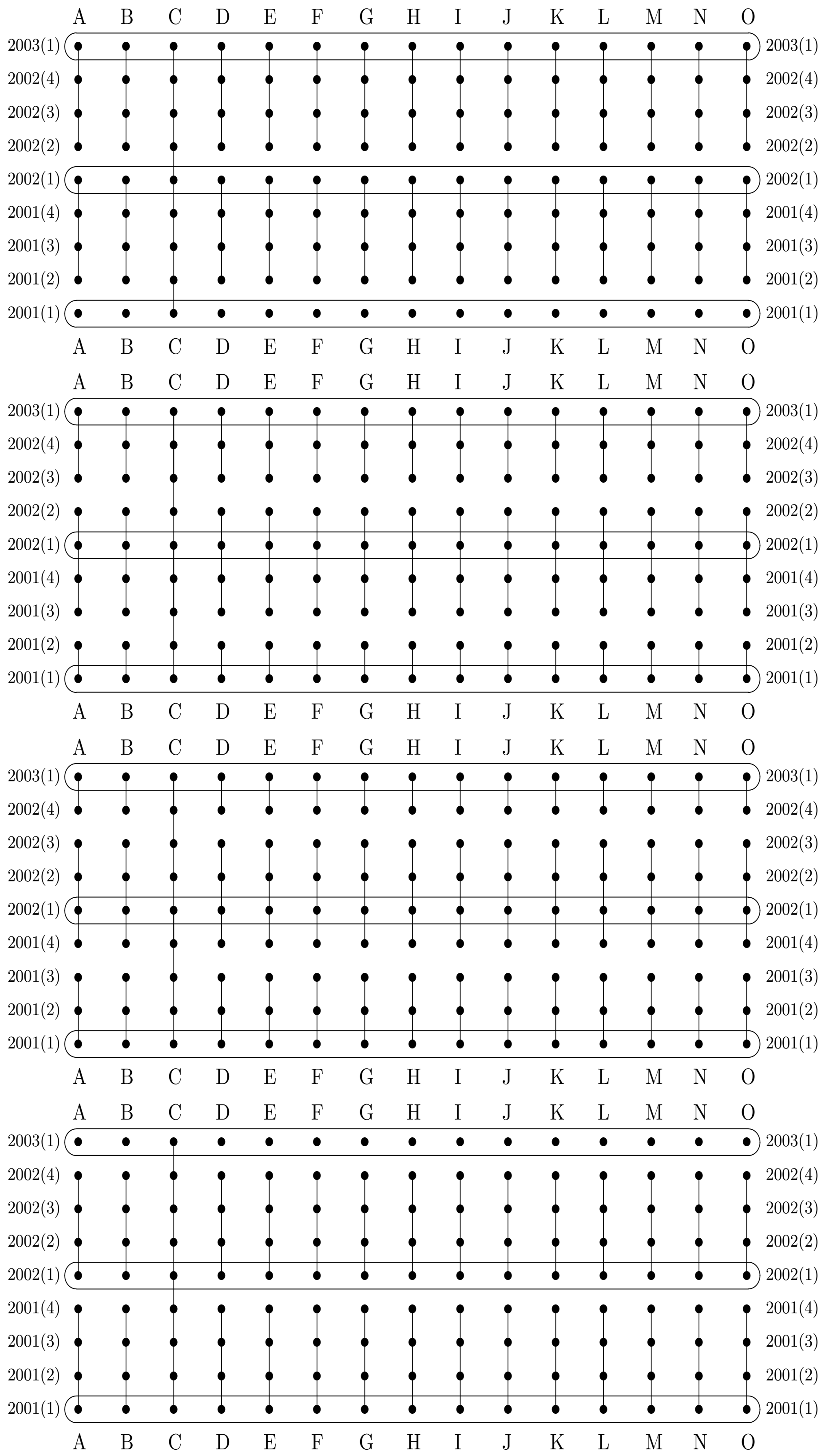


Figure 6: House Price Indexes for Sydney by Region

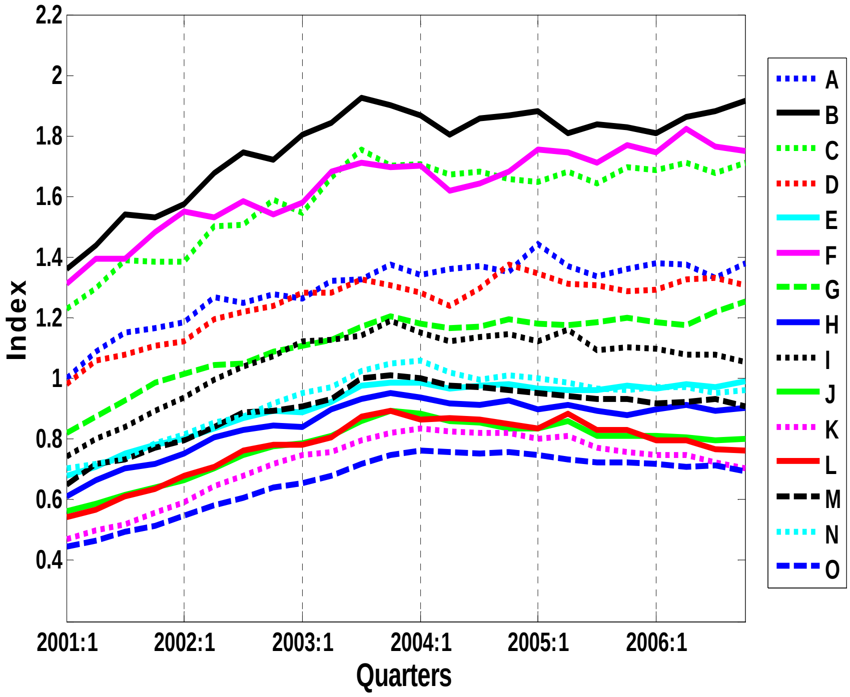


Figure 7: Temporal Price Indexes for Sydney

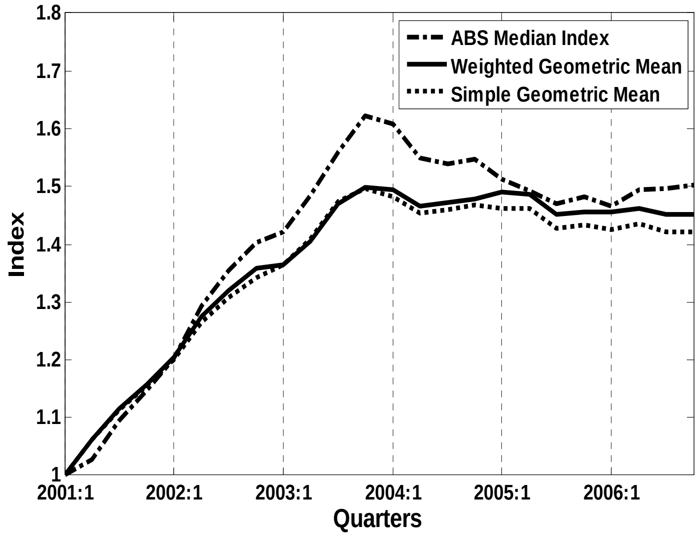
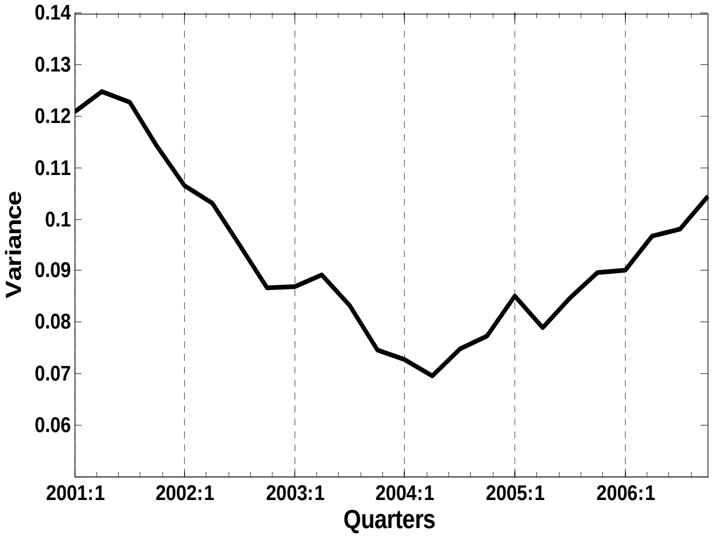


Figure 8: σ-Convergence Results



Appendix table 1: Estimated Coefficients from Equation (7) in Six Periods

Variables		Period1	Period2	Period3	Period4	Period5	Period6
Intercept		9.0987 ⊥	9.5678	9.7711	9.9082	9.8815	10.2320
Time	**Qtr2**	0.0843	0.0675	0.0441	0.0157	-0.0498	-0.0038
Dummies	**Qtr3**	0.1413	0.0501	0.0484	0.0241	-0.0745	-0.0360
(Base:Qtr1)	**Qtr4**	0.1515	0.0770	0.0830	0.0081	-0.0558	-0.0006
	Qtr5	0.1701	0.0644	0.0578	0.0729	-0.0423	-
Regions	**B**	-0.2681	-0.2400	-0.2088	-0.1299	-0.3351	-0.3895
(Base:Region A)	**C**	-0.0543 ⊥	-0.0698	-0.0622	-0.0179	-0.1219	-0.0260
	D	-0.0253 ⊥	-0.0260 ⊥	-0.0484	-0.0064	-0.0566	-0.0476
	E	-0.1037	-0.1680	-0.1763	-0.0890	-0.2382	-0.1707
	F	0.1869	0.1346	0.1701	0.2099	0.0840	0.0814
	G	0.1201	0.1354	0.1881	0.2350	0.1744	0.2514
	H	-0.4266	-0.3580	-0.3970	-0.3417	-0.4530	-0.4869
	I	-0.5445	-0.4913	-0.4360	-0.4065	-0.5786	-0.6355
	J	-0.4552	-0.5425	-0.5171	-0.3134	-0.4915	-0.4899
	K	-0.5982	-0.6695	-0.5749	-0.5258	-0.6744	-0.6836
	L	-0.3826	-0.4230	-0.3679	-0.3086	-0.4955	-0.5318
	M	-0.2606	-0.2726	-0.2389	-0.2150	-0.3548	-0.3537
	N	-0.3559	-0.4088	-0.3585	-0.3230	-0.4413	-0.4383
	O	-0.6800	-0.7278	-0.6964	-0.6567	-0.7530	-0.7599
Time- Region	**Qtr2-B**	-0.0253 ⊥	-0.0067 ⊥	-0.0236	-0.0510	0.0103	0.0323
Inter. Dummies	**Qtr2-C**	-0.0283 ⊥	0.0156 ⊥	0.0282	-0.0359	0.0730	0.0198
	Qtr2-D	-0.0068 ⊥	-0.0057 ⊥	-0.0437	-0.0503	0.0241	0.0276
	Qtr2-E	-0.0427	-0.0312 ⊥	-0.0090	-0.0348	0.0459	0.0183
	Qtr2-F	-0.0242 ⊥	-0.0817	0.0165	-0.0633	0.0448	0.0472
	Qtr2-G	-0.0217 ⊥	-0.0396	-0.0250	-0.0259	0.0458	-0.0050
	Qtr2-H	0.0007 ⊥	-0.0013 ⊥	0.0222	-0.0398	0.0645	0.0160
	Qtr2-I	-0.0105 ⊥	-0.0083 ⊥	-0.0406	-0.0444	0.0806	-0.0118
	Qtr2-J	-0.0460	-0.0132 ⊥	-0.0152	-0.0432	0.0768	-0.0044
	Qtr2-K	-0.0257 ⊥	0.0181 ⊥	-0.0298	-0.0254	0.0606	0.0055
	Qtr2-L	-0.0387	-0.0246 ⊥	-0.0158	-0.0054	0.1088	0.0027
	Qtr2-M	0.0107 ⊥	-0.0129 ⊥	-0.0145	-0.0388	0.0377	0.0082
	Qtr2-N	-0.0673	-0.0180 ⊥	-0.0239	-0.0536	0.0322	0.0045
	Qtr2-O	-0.0458	-0.0096 ⊥	-0.0073	-0.0229	0.0349	-0.0106
	Qtr3-B	-0.0150 ⊥	0.0525	0.0145	-0.0299	0.0494	0.0753
	Qtr3-C	-0.0181 ⊥	0.0357	0.0774	-0.0360	0.0737	0.0321

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

Qtr3-D	-0.0485	0.0305 ⊥	-0.0142	-0.0151	0.0432	0.0655
Qtr3-E	-0.0357	0.0269 ⊥	0.0434	-0.0313	0.0701	0.0415
Qtr3-F	-0.0795	-0.0301 ⊥	0.0302	-0.0577	0.0501	0.0490
Qtr3-G	-0.0199 ⊥	-0.0156 ⊥	0.0083	-0.0322	0.0757	0.0636
Qtr3-H	-0.0031 ⊥	0.0481	0.0541	-0.0511	0.0690	0.0291
Qtr3-I	-0.0191 ⊥	0.0525	-0.0299	-0.0385	0.0465	0.0202
Qtr3-J	-0.0579	0.0641	0.0416	-0.0605	0.0433	0.0164
Qtr3-K	-0.0430	0.0916	0.0164	-0.0367	0.0360	0.0037
Qtr3-L	-0.0223 ⊥	0.0636	0.0597	-0.0197	0.0690	-0.0045
Qtr3-M	-0.0260 ⊥	0.0635	0.0515	-0.0540	0.0516	0.0476
Qtr3-N	-0.0936	0.0194 ⊥	0.0270	-0.0869	0.0400	0.0169
Qtr3-O	-0.0437	0.0491	0.0432	-0.0358	0.0438	0.0310
Qtr4-B	-0.0303	0.0094 ⊥	-0.0316	-0.0092	0.0244	0.0573
Qtr4-C	-0.0305 ⊥	0.0613	0.0113	-0.0374	0.0855	0.0154
Qtr4-D	-0.0283 ⊥	0.0202 ⊥	-0.0631	0.0600	0.0123	0.0111
Qtr4-E	-0.0121 ⊥	0.0309 ⊥	0.0189	-0.0119	0.0662	0.0246
Qtr4-F	-0.0310 ⊥	-0.0850	-0.0119	-0.0182	0.0655	0.0055
Qtr4-G	0.0308 ⊥	-0.0072 ⊥	0.0040	0.0068	0.0712	0.0565
Qtr4-H	0.0088 ⊥	0.0376	0.0407	-0.0182	0.0314	0.0029
Qtr4-I	0.0298	0.0579	-0.0218	-0.0143	0.0352	-0.0413
Qtr4-J	-0.0265 ⊥	0.0767	0.0456	-0.0654	0.0290	-0.0109
Qtr4-K	0.0178 ⊥	0.1188	0.0092	-0.0242	-0.0009 ⊥	-0.0621
Qtr4-L	0.0096 ⊥	0.0650	0.0488	-0.0200	0.0490	-0.0437
Qtr4-M	0.0149 ⊥	0.0379	0.0263	-0.0460	0.0354	-0.0134
Qtr4-N	-0.0451	0.0414	0.0147	-0.0577	0.0148	-0.0097
Qtr4-O	-0.0088 ⊥	0.0774	0.0472	-0.0132	0.0239	-0.0369
Qtr5-B	-0.0207 ⊥	0.0710	-0.0236	-0.0650	0.0012 ⊥	-
Qtr5-C	-0.0510	0.0462	0.0396	-0.1082	0.0657	-
Qtr5-D	-0.0338	0.0668	-0.0560	-0.0256	0.0030	-
Qtr5-E	-0.0013 ⊥	0.0387	0.0416	-0.0917	0.0431	-
Qtr5-F	-0.0024 ⊥	-0.0461 ⊥	0.0142	-0.0407	0.0360	-
Qtr5-G	0.0408	0.0217 ⊥	0.0065	-0.0704	0.0443	-
Qtr5-H	0.0386	0.0460	0.0505	-0.1148	0.0422	-
Qtr5-I	0.0605	0.1144	-0.0300	-0.0983	0.0177	-
Qtr5-J	-0.0043 ⊥	0.1011	0.0595	-0.1314	0.0158	-
Qtr5-K	0.0555	0.1718	0.0494	-0.1098	-0.0280	-
Qtr5-L	0.0535	0.0800	0.0377	-0.1062	-0.0036	-
Qtr5-M	0.0275 ⊥	0.0668	0.0412	-0.1206	0.0060	-
Qtr5-N	-0.0252 ⊥	0.0905	0.0496	-0.1278	0.0112	-
Qtr5-O	0.0340	0.1121	0.0918	-0.0922	0.0058	-

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

Postcodes	**2000**	0.1961	0.1747	0.1640	0.1074	0.1252	0.1576
	2007	-0.0141 ⊥	0.0854	-0.0272	-0.1382	-0.0922	-0.1884
	2008	-0.1182	-0.1010	-0.1001	-0.1101	-0.0744	-0.1275
	2009	0.0861	0.0890	0.0403	-0.0157	-0.0137	0.0379
	2011	0.0070 ⊥	0.0603	0.0845	0.0667	0.1013	0.1621
	2015	-0.0865	-0.0310	-0.0443	-0.1294	-0.1351	-0.1195
	2016	-0.1423	-0.1287	-0.0997	-0.1832	-0.1237	-0.1437
	2017	-0.0471	-0.0348	-0.0275	-0.0664	-0.0861	-0.1042
	2018	-0.2213	-0.2029	-0.1992	-0.3169	-0.2653	-0.3047
	2019	-0.2177	-0.1795	-0.1650	-0.1826	-0.1968	-0.2115
	2020	-0.2723	-0.2795	-0.1552	-0.2144	-0.1537	-0.1516
	2021	0.4453	0.4075	0.3663	0.3639	0.4967	0.5370
	2022	0.3789	0.3625	0.3328	0.3655	0.4269	0.4443
	2023	0.4498	0.4954	0.4585	0.4783	0.5524	0.5442
	2024	0.3730	0.3685	0.4052	0.3710	0.4686	0.5093
	2025	0.5686	0.5083	0.4813	0.4618	0.5587	0.5843
	2026	0.4155	0.3751	0.4167	0.4372	0.4720	0.5192
	2027	0.6120	0.6173	0.5688	0.7010	0.7011	0.7246
	2028	0.5434	0.5224	0.5406	0.5435	0.5699	0.6158
	2029	0.4866	0.4758	0.4318	0.4450	0.4853	0.5352
	2030	0.4319	0.4018	0.3710	0.4314	0.4555	0.5125
	2031	0.3326	0.3408	0.3289	0.3313	0.3898	0.4481
	2032	0.1900	0.1745	0.1432	0.1189	0.1798	0.2113
	2033	0.1950	0.2769	0.2415	0.2759	0.2815	0.3236
	2034	0.4218	0.3896	0.3843	0.3937	0.4468	0.4929
	2035	0.2181	0.2418	0.2316	0.2365	0.2618	0.2469
	2037	0.0404 ⊥	-0.0111 ⊥	0.0667	0.0091	-0.0356	-0.0720
	2038	-0.0123 ⊥	-0.0381 ⊥	-0.0562	-0.0133	-0.0370	-0.0442
	2039	0.0633	0.0143 ⊥	0.0148	0.0228	0.0212	0.0281
	2040	-0.0939	-0.0916	-0.0446	-0.0664	-0.0834	-0.1347
	2041	0.1366	0.1374	0.1245	0.1373	0.1196	0.0804
	2042	-0.1292	-0.1329	-0.1531	-0.1260	-0.1235	-0.1810
	2043	-0.0654	-0.0854	-0.0647	-0.0383	-0.0443	-0.1106
	2044	-0.2946	-0.3196	-0.3010	-0.2541	-0.2917	-0.3274
	2045	-0.2291	-0.1687	-0.1522	-0.1872	-0.1117	-0.0944
	2046	-0.0899	-0.0095 ⊥	-0.0069	-0.0828	-0.0566	-0.0846
	2047	0.0276 ⊥	0.0537	0.0731	0.0445	0.0240	0.0488
	2048	-0.1584	-0.1284	-0.1141	-0.1154	-0.1332	-0.1567
	2049	-0.2102	-0.1592	-0.1224	-0.1490	-0.1762	-0.2182
	2060	0.3311	0.2570	0.3067	0.2713	0.2583	0.2889

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

2061	0.5000	0.4906	0.5616	0.4987	0.5062	0.4394
2062	0.1910	0.2178	0.2457	0.2315	0.1808	0.2010
2063	0.1925	0.2171	0.2475	0.1839	0.1678	0.1948
2064	0.1151	0.1192	0.1145	0.1056	0.1143	0.1013
2065	0.1640	0.1849	0.2237	0.2014	0.1506	0.1409
2066	0.0122 ⊥	0.0104 ⊥	0.0652	0.0392	0.0000 ⊥	-0.0063
2067	0.0435	0.0925	0.1285	0.1052	0.0748	0.0477
2068	0.0527	0.0715	0.0909	0.0459	0.0571	0.0570
2070	0.1609	0.1464	0.1787	0.1826	0.1972	0.2214
2071	0.1948	0.1711	0.1514	0.1739	0.1796	0.1601
2072	0.1718	0.1473	0.1323	0.1057	0.1344	0.1711
2073	0.0549	0.0394 ⊥	0.0817	0.0423	0.0729	0.0881
2074	0.0189 ⊥	0.0309 ⊥	0.0307	0.0407	0.0328	0.0622
2075	0.0006 ⊥	0.0098 ⊥	0.0361	0.0292	0.0563	0.0374
2076	-0.0355 ⊥	-0.0358 ⊥	-0.0411	-0.0508	-0.0449	-0.0311
2077	-0.1016	-0.1101	-0.1292	-0.1494	-0.1253	-0.1698
2079	-0.2172	-0.2619	-0.2400	-0.2788	-0.2244	-0.2759
2080	-0.2969	-0.4438	-0.2860	-0.2984	-0.2946	-0.3273
2081	-0.2395	-0.2752	-0.2512	-0.1998	-0.2206	-0.2484
2082	-0.2628	-0.2858	-0.2512	-0.2531	-0.2542	-0.2874
2083	-0.1186	-0.0785 ⊥	-0.1344	-0.2094	-0.1087	-0.1492
2084	-0.0308 ⊥	0.0672 ⊥	0.0690	0.0245	0.0112	0.0533
2085	-0.0342 ⊥	-0.0888	-0.0211	-0.0331	-0.0602	-0.0620
2086	-0.0156 ⊥	-0.0031 ⊥	0.0064	-0.0231	-0.0124	-0.0319
2088	0.0058 ⊥	0.0272 ⊥	0.0290	0.0160	0.0473	0.0342
2089	-0.0206 ⊥	-0.0144 ⊥	0.0055	-0.0442	-0.0399	-0.0318
2092	-0.0358 ⊥	-0.0040 ⊥	-0.0293	-0.1252	-0.1680	-0.1616
2094	0.0260 ⊥	0.0266 ⊥	0.0588	-0.0103	-0.0492	-0.0326
2095	0.1197	0.1644	0.1636	0.0808	0.0758	0.0798
2096	-0.0902	-0.0435 ⊥	-0.0557	-0.1014	-0.1423	-0.1804
2097	-0.1524	-0.1179	-0.1716	-0.2313	-0.2599	-0.2364
2099	-0.1929	-0.1760	-0.2151	-0.2726	-0.3171	-0.3358
2100	-0.2862	-0.2710	-0.3176	-0.3665	-0.3779	-0.3908
2101	-0.2014	-0.1350	-0.1704	-0.2494	-0.2859	-0.2828
2102	-0.2924	-0.2849	-0.3800	-0.3457	-0.3916	-0.5319
2103	-0.1807	-0.1447	-0.1919	-0.1899	-0.2466	-0.2614
2106	-0.1642	-0.1565	-0.1354	-0.2229	-0.2299	-0.2687
2107	-0.1970	-0.1727	-0.2191	-0.2483	-0.2762	-0.2837
2110	0.5276	0.5222	0.5400	0.5584	0.5655	0.5941
2111	0.3773	0.3198	0.3011	0.3436	0.3772	0.3810

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

2112	0.1988	0.1767	0.1960	0.1669	0.1945	0.2001
2113	0.1812	0.2032	0.1798	0.1794	0.2116	0.1999
2114	0.0879	0.1592	0.1892	0.1458	0.1899	0.1763
2115	-0.0160 ⊥	-0.0328 ⊥	0.0005 ⊥	0.0101	0.0435	0.0069
2116	0.0363 ⊥	0.1345	0.3093	0.0028 ⊥	-0.0016 ⊥	0.0087 ⊥
2117	0.0265 ⊥	0.0302 ⊥	0.0638	0.0663	0.0555	0.0251
2118	0.0308 ⊥	0.0345 ⊥	0.0498	0.0574	0.0325	0.0049
2119	0.1663	0.1815	0.1813	0.2012	0.1937	0.1792
2120	0.0550	0.0395	0.0830	0.0718	0.0565	0.0324
2121	0.1748	0.1891	0.2067	0.2007	0.1911	0.2008
2125	0.0395	0.0607	0.0905	0.0973	0.0573	0.0246
2127	0.3585	0.3308	0.3107	0.3015	0.2815	0.3809
2128	-0.0222 ⊥	0.0506 ⊥	0.0748	0.0943	0.1339	0.1549
2130	0.2863	0.2543	0.2967	0.2646	0.2868	0.4118
2131	0.2423	0.2035	0.2228	0.1905	0.2011	0.3138
2132	0.3271	0.2783	0.3091	0.2685	0.3028	0.4004
2133	0.1870	0.1620	0.1732	0.1808	0.1744	0.2558
2134	0.3686	0.3331	0.2907	0.2622	0.2817	0.3668
2135	0.3682	0.3484	0.3488	0.3133	0.3081	0.4123
2136	0.2309	0.1969	0.2245	0.2185	0.1876	0.2352
2137	0.4605	0.4130	0.4167	0.4095	0.3763	0.4825
2138	0.3273	0.3055	0.3641	0.3169	0.2885	0.3490
2139	0.0633	0.0143 ⊥	0.0148	0.0228	0.0212	0.0281
2140	0.2286	0.2582	0.2387	0.2104	0.1813	0.2724
2141	0.0479	0.0679	0.0441	0.0343	0.0328	0.0460
2142	-0.0455	-0.0564	-0.0026	-0.0217	-0.0295	-0.0231
2143	-0.0890	-0.0346 ⊥	-0.0794	-0.0960	-0.0985	-0.0239
2144	-0.0513	0.0183 ⊥	0.0037	0.0058	-0.0147	0.0217
2146	-0.0589 ⊥	-0.0167 ⊥	-0.0157	-0.1167	-0.0734	-0.1113
2147	-0.1116	-0.0706 ⊥	-0.0785	-0.1851	-0.1566	-0.1757
2148	-0.1467	-0.1099	-0.0744	-0.1978	-0.1790	-0.1867
2150	0.0498 ⊥	0.1044	0.0823	-0.0530	-0.0313	-0.0329
2151	0.0813 ⊥	0.1040	0.0989	-0.0340	-0.0015 ⊥	0.0007 ⊥
2152	0.0823 ⊥	0.1068	0.1129	-0.0439	-0.0354	-0.0067 ⊥
2153	0.1027	0.1204	0.1278	0.0066 ⊥	0.0208	0.0194
2154	0.1566	0.1905	0.1793	0.0472	0.0620	0.0762
2156	0.1591	0.1719	0.1639	0.0472	0.0429	0.0160
2157	-0.0830 ⊥	-0.0673 ⊥	-0.0181 ⊥	-0.1374	-0.0868	-0.1225
2158	0.1751	0.1879	0.1866	0.0475	0.0508	0.0824

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

2160	0.1868	0.1790	0.1587	0.1498	0.1337	0.1639
2161	0.0589	0.0810	0.1265	0.1159	0.1128	0.0808
2162	0.0808	0.1066	0.1088	0.0679	0.0871	0.1229
2163	-0.0849	-0.0846	-0.0802	-0.0777	-0.0415	-0.0785
2164	-0.0272 ⊥	0.0142 ⊥	0.0306	0.0391	0.0214	0.0379
2165	0.0180 ⊥	0.0466	0.0489	0.0632	0.0665	0.0456
2166	-0.1135	-0.0812	-0.0551	-0.0374	-0.0234	-0.0152
2167	-0.0775	-0.0690	-0.0573	-0.0627	0.0226	-0.0087
2168	-0.0574	-0.0379 ⊥	-0.0129	-0.0129	-0.0332	-0.0552
2170	0.0523	0.0431	0.0469	0.0558	0.0497	0.0544
2171	0.1033	0.0873	0.1224	0.1055	0.0886	0.1009
2176	0.0585	0.0454 ⊥	0.0477	0.0228	0.0423	0.0734
2190	-0.0436	-0.0330	-0.0180	-0.0291	-0.0330	-0.0492
2191	0.1347	0.1145	0.1117	0.1238	0.1256	0.1699
2192	-0.0191 ⊥	0.0793	0.0517	0.0112	-0.0122	0.0331
2195	-0.1445	-0.1037	-0.0821	-0.1292	-0.1285	-0.1227
2196	-0.0628	-0.0820	-0.0435	-0.0445	-0.0342	-0.0503
2197	-0.1288	-0.1227	-0.0996	-0.0910	-0.0915	-0.0504
2198	-0.0129 ⊥	-0.0957	-0.0859	-0.0495	-0.0277	-0.0425
2199	-0.0847	-0.1024	-0.0857	-0.0778	0.0144	-0.0716
2203	-0.0022 ⊥	0.0443	0.0573	0.0707	0.0576	0.0771
2204	-0.0172 ⊥	0.0139 ⊥	0.0390	0.0505	0.0944	0.0505
2205	0.0197 ⊥	0.0829	0.0592	0.0718	0.0262	0.0055
2206	0.0076 ⊥	0.0233 ⊥	0.0190	0.0118	0.0184	0.0040
2207	-0.0556	-0.0349	-0.0445	-0.0031	-0.0215	-0.0390
2208	-0.0816	-0.0099 ⊥	-0.0107	-0.0394	0.0237	-0.0120
2209	-0.1186	-0.0990	-0.0744	-0.0707	-0.1027	-0.1158
2210	-0.1232	-0.1013	-0.0951	-0.0403	-0.0979	-0.1146
2211	-0.2279	-0.1614	-0.1448	-0.1247	-0.1363	-0.1545
2212	-0.2274	-0.1975	-0.2050	-0.1801	-0.1916	-0.2125
2213	-0.2079	-0.1774	-0.1777	-0.1539	-0.1444	-0.1722
2214	-0.1704	-0.2085	-0.2490	-0.2573	-0.2077	-0.2298
2216	0.0776	0.0824	0.0880	0.1156	0.0762	0.0774
2217	0.0595	0.0695	0.0621	0.0831	0.0499	0.0390
2218	-0.0356	0.0041 ⊥	-0.0030	0.0026	-0.0024	-0.0011 ⊥
2224	0.1994	0.2455	0.2892	0.2553	0.2854	0.3322
2225	0.0436 ⊥	0.1049	0.2014	0.1598	0.1344	0.1344
2226	0.0389	0.0566	0.0622	0.0592	0.0460	0.0641

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

	2227	0.0841	0.1474	0.1480	0.1345	0.1691	0.1231
	2228	0.1102	0.1368	0.1401	0.1501	0.1348	0.1212
	2229	0.1921	0.2332	0.2375	0.2110	0.2329	0.2410
	2230	0.2749	0.2955	0.3299	0.3397	0.3339	0.3558
	2231	-0.0797	-0.1477	-0.0165	-0.0468	-0.0383	-0.0281
	2232	0.0591	0.0563	0.0696	0.0852	0.0790	0.0897
	2233	-0.0584	-0.0402	-0.0295	-0.0289	-0.0173	-0.0296
	2745	0.1398	0.1481	0.1451	0.1566	0.1601	0.1259
	2747	-0.0680	-0.0314 ⊥	-0.0162	0.0035	-0.0179	-0.0217
	2750	0.0085 ⊥	0.0260 ⊥	0.0317	0.0551	0.0520	0.0557
	2760	-0.0952	-0.0641	-0.0493	-0.0232	-0.0506	-0.0664
	2761	0.0327 ⊥	0.0382	0.0571	0.1012	0.0856	0.0632
	2763	0.1434	0.1775	0.1558	0.1739	0.1434	0.1494
	2766	-0.0228 ⊥	0.0214 ⊥	0.0464	0.0971	0.1226	0.1146
	2767	0.0588	0.0927	0.0579	0.0891	0.0954	0.0368
	2768	0.2592	0.2746	0.2621	0.2693	0.2601	0.2454
	2770	-0.1824	-0.1159	-0.0869	-0.0382	-0.0854	-0.1340
	2773	0.1737	0.1850	0.1453	0.2052	0.1592	0.1441
	2774	0.0533	0.0352 ⊥	0.0313	0.0835	0.1126	0.0813
Dwelling Type (Base: House)	**Units**	-0.1841	-0.2162	-0.2604	-0.1924	-0.2358	-0.2774
Bedroom Counts (Base: 2)	**1**	-0.1474	-0.1854	-0.1506	-0.1397	-0.1810	-0.1959
	3	0.1230	0.1161	0.1473	0.1752	0.1625	0.1572
	4	0.2443	0.2285	0.2569	0.2192	0.1995	0.2094
	5	0.1978	0.1855	0.2024	0.2514	0.1784	0.3206
Bathroom Counts (Base: 1)	**2**	0.0887	0.0961	0.1221	0.1405	0.1559	0.1508
	3	0.3154	0.3429	0.3828	0.4590	0.4907	0.3991
	4	0.6079	0.3057	0.4903	0.5263	0.5983	0.5554
Lot Size ($meter^2/10^3$)		0.4328	0.3768	0.3974	0.4321	0.3117	0.2686
Lot Size^2 ($meter^4/10^6$)		-0.0325	-0.0521	-0.0538	-0.0640	-0.0514	-0.0492

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

Bedroom-Region Inter. Dummies	**1-B**	-0.0622	-0.0438	-0.0198	-0.0150	-0.0267	-0.0544
	3-B	0.0469	0.0380	0.0347	0.0137	0.0195	0.0219
	4-B	0.0317 ⊥	0.0136 ⊥	0.0275	0.0292	0.0173	0.0784
	5-B	0.0561 ⊥	0.0715 ⊥	0.0875	-0.0181	-0.0078 ⊥	-0.0529
	1-C	0.0100 ⊥	0.0168 ⊥	0.0060	-0.0074	0.0255	0.0111
	3-C	-0.0137 ⊥	0.0008 ⊥	-0.0313	-0.0269	-0.0055	-0.0186
	4-C	-0.0343 ⊥	-0.0050 ⊥	-0.0396	-0.0065	0.0544	0.0553
	5-C	0.0592 ⊥	0.1112 ⊥	0.0337	-0.0231	0.0322	0.0271
	1-D	0.0108 ⊥	0.0012 ⊥	-0.0110	-0.0147	-0.0068	-0.0254
	3-D	-0.0025 ⊥	-0.0079 ⊥	-0.0142	-0.0411	-0.0310	-0.0182
	4-D	-0.0503 ⊥	-0.0271 ⊥	-0.0003 ⊥	-0.0278	0.0380	0.0145
	5-D	-0.0099 ⊥	0.0713 ⊥	0.0696	-0.0413	0.0440	-0.0062
	1-E	0.1204	0.1073	0.1220	0.1077	0.1439	0.0950
	3-E	-0.0958	-0.0745	-0.1088	-0.1214	-0.0905	-0.1160
	4-E	-0.1656	-0.1418	-0.1460	-0.1473	-0.0866	-0.1374
	5-E	-0.1342	-0.0541 ⊥	-0.0363	-0.1255	-0.0817	-0.1848
	1-F	-0.0385 ⊥	-0.0507	-0.0738	-0.0645	-0.0680	-0.1180
	3-F	0.1120	0.0990	0.0813	0.0665	0.0512	0.0272
	4-F	0.1776	0.1609	0.1945	0.1400	0.1641	0.0925
	5-F	0.0726 ⊥	0.2553	0.2760	0.1693	0.0762	-0.0201
	1-G	-0.0152 ⊥	-0.0309 ⊥	-0.0284	-0.0409	0.0042	-0.0237
	3-G	-0.0182 ⊥	-0.0621	-0.0488	-0.0701	-0.0519	-0.0585
	4-G	-0.1067	-0.1307	-0.1015	-0.1095	-0.0944	-0.0887
	5-G	-0.0605 ⊥	-0.0536 ⊥	-0.0496	-0.1397	-0.1035	-0.1750
	1-H	0.0194 ⊥	0.0009 ⊥	0.0027 ⊥	0.0043 ⊥	0.0090	-0.0409
	3-H	-0.0219 ⊥	-0.0494	-0.0519	-0.0712	-0.0515	-0.0458
	4-H	-0.0997	-0.1292	-0.1151	-0.1222	-0.0507	-0.0817
	5-H	-0.0674 ⊥	-0.0504 ⊥	-0.0522	-0.1379	-0.0505	-0.1588
	1-I	0.0594	0.0716	0.0398	0.0763	0.0885	0.0648
	3-I	-0.0906	-0.0882	-0.1058	-0.1370	-0.0967	-0.0840
	4-I	-0.1603	-0.1581	-0.1417	-0.1183	-0.0484	-0.0919

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

	5-I	-0.0161 ⊥	0.0379 ⊥	0.0730	-0.0183	0.0760	-0.1100
	1-J	0.0895	0.1109	0.1167	0.1704	0.2385	0.1824
	3-J	-0.0943	-0.0991	-0.1404	-0.1422	-0.1174	-0.1252
	4-J	-0.2079	-0.1937	-0.2135	-0.1957	-0.1233	-0.1330
	5-J	-0.1368	-0.0933 ⊥	-0.0871	-0.1470	-0.1017	-0.1977
	1-K	0.1027	0.0871	0.1226	0.1024	0.1767	0.1357
	3-K	-0.0937	-0.0883	-0.1266	-0.1415	-0.0987	-0.1195
	4-K	-0.1826	-0.1490	-0.1675	-0.1530	-0.0809	-0.1063
	5-K	-0.1008	-0.0606 ⊥	-0.0583	-0.1258	-0.0443	-0.1641
	1-L	0.1166	0.1511	0.1545	0.2018	0.1856	0.1617
	3-L	-0.0792	-0.0953	-0.1153	-0.1241	-0.1138	-0.0947
	4-L	-0.1805	-0.1751	-0.1738	-0.1725	-0.1258	-0.1292
	5-L	-0.0663 ⊥	-0.1101	-0.0387	-0.0650	-0.0269	-0.0973
	1-M	0.0586	0.0522	0.0725	0.0689	0.0928	0.0819
	3-M	-0.0543	-0.0707	-0.0871	-0.0920	-0.0618	-0.0727
	4-M	-0.1451	-0.1515	-0.1262	-0.0927	-0.0717	-0.1030
	5-M	-0.1143	-0.0733 ⊥	-0.0009 ⊥	-0.0205	-0.0011 ⊥	-0.0855
	1-N	-0.0173 ⊥	-0.0470 ⊥	-0.0343	0.0069	-0.0033	-0.0724
	3-N	-0.0600	-0.0509	-0.0918	-0.0929	-0.0790	-0.0940
	4-N	-0.1095	-0.1034	-0.1308	-0.1268	-0.0772	-0.1135
	5-N	-0.0812 ⊥	-0.0271	-0.0495	-0.0969	-0.0281	-0.2107
	1-O	0.0938	0.0937	0.1243	0.1814	0.1416	0.1424
	3-O	-0.0747	-0.1106	-0.1361	-0.1442	-0.1350	-0.1410
	4-O	-0.1811	-0.2071	-0.2057	-0.1876	-0.1321	-0.1584
	5-O	-0.1459	-0.1400	-0.1338	-0.1857	-0.1158	-0.2300
Bathroom-	**2-B**	0.0219 ⊥	0.0283 ⊥	0.0148	0.0039	-0.0054	0.0053
Region	**3-B**	-0.0380 ⊥	-0.0320 ⊥	-0.0961	-0.0921	-0.1035	-0.0227
Inter. Dummies	**4-B**	-0.2748	0.0275 ⊥	-0.1530	-0.1331	-0.1224	-0.1132

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

2-C	-0.0163 ⊥	-0.0158 ⊥	-0.0451	-0.0486	-0.0262	-0.0080
3-C	-0.0563 ⊥	-0.0797	-0.1278	-0.0807	-0.1804	-0.0473
4-C	-0.1278 ⊥	0.1082 ⊥	-0.1379	0.0100 ⊥	-0.0670	-0.1868
2-D	-0.0071 ⊥	-0.0308	-0.0472	-0.0539	-0.0703	-0.0591
3-D	-0.1335	-0.2067	-0.1511	-0.1954	-0.2276	-0.1564
4-D	-0.2509	-0.1609 ⊥	-0.1491	-0.0987	-0.3145	-0.3222
2-E	-0.0798	-0.0887	-0.1107	-0.1145	-0.1111	-0.1125
3-E	-0.3027	-0.3284	-0.3435	-0.3377	-0.3348	-0.2590
4-E	-0.3796	-0.1562 ⊥	-0.2585	-0.2153	-0.2068	-0.1637
2-F	0.0415 ⊥	0.0230 ⊥	0.0124	-0.0478	0.0053	-0.0268
3-F	-0.0521 ⊥	-0.1702	-0.1820	-0.1445	-0.1033	-0.0962
4-F	-0.1821 ⊥	-0.2235	-0.4056	-0.2825	-0.2463	-0.3172
2-G	-0.0167 ⊥	-0.0207 ⊥	-0.0431	-0.0424	-0.0201	-0.0267
3-G	-0.1603	-0.1923	-0.1778	-0.1783	-0.1739	-0.1106
4-G	-0.0797 ⊥	0.0386 ⊥	-0.0856	0.1099	-0.0722	-0.0273 ⊥
2-H	-0.0617	-0.0673	-0.0945	-0.1078	-0.0857	-0.0875
3-H	-0.2741	-0.3009	-0.3389	-0.3497	-0.3012	-0.2431
4-H	-0.4217	-0.1023 ⊥	-0.2426	-0.1773	-0.1695	-0.1296
2-I	-0.0649	-0.0709	-0.1182	-0.1169	-0.1084	-0.1046
3-I	-0.2643	-0.2416	-0.2434	-0.3129	-0.2343	-0.2182
4-I	-0.4198	0.1966 ⊥	0.0553	-0.1809	0.0108 ⊥	-0.0478
2-J	-0.0773	-0.0878	-0.1192	-0.1238	-0.1261	-0.1295
3-J	-0.2878	-0.3099	-0.3586	-0.3530	-0.3332	-0.2749
4-J	-0.2236 ⊥	-0.1857 ⊥	-0.2583	-0.3089	-0.1738	-0.1641
2-K	-0.0898	-0.0975	-0.1270	-0.1440	-0.1260	-0.1091
3-K	-0.2952	-0.2570	-0.2954	-0.3232	-0.2689	-0.2741
4-K	0.0226 ⊥	0.3595	-0.1498	0.0435 ⊥	0.0554	-0.0068 ⊥
2-L	-0.0900	-0.1055	-0.1275	-0.1468	-0.1404	-0.1192
3-L	-0.2572	-0.3149	-0.3154	-0.3496	-0.3038	-0.2363
4-L	-0.3719 ⊥	-0.4329 ⊥	-0.2769	0.1604	0.4015 ⊥	-0.2341
2-M	-0.0737	-0.0816	-0.1093	-0.1227	-0.1136	-0.1105

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

	3-M	-0.2896	-0.2905	-0.3149	-0.3320	-0.3440	-0.2356
	4-M	-0.4048	-0.1304 ⊥	-0.2247	-0.2578	-0.3488	-0.2108
	2-N	-0.0636	-0.0769	-0.1019	-0.1179	-0.1050	-0.0865
	3-N	-0.2573	-0.2710	-0.2994	-0.2965	-0.2704	-0.1726
	4-N	-0.2023 ⊥	0.0493 ⊥	-0.1268	-0.1907	-0.0512	-0.0605
	2-O	-0.0961	-0.1071	-0.1300	-0.1405	-0.1196	-0.1014
	3-O	-0.3258	-0.3634	-0.3808	-0.4099	-0.3702	-0.2797
	4-O	-0.3491 ⊥	0.0305 ⊥	0.0530	-0.2907	-0.1034 ⊥	-0.0948
Dwelling type- Region Inter. Dummies	**Unit-B**	-0.0418	-0.0810	-0.0353	-0.1905	-0.0715	-0.0445
	Unit-C	0.0932	0.0639	0.0952	0.0262	0.0282	0.0438
	Unit-D	-0.0864	-0.0928	-0.0604	-0.1712	-0.0975	-0.1017
	Unit-E	0.0775	0.1158	0.1534	0.0491	0.0892	0.0782
	Unit-F	-0.0949	-0.0060 ⊥	-0.0773	-0.1067	-0.0497	0.0344
	Unit-G	-0.0053 ⊥	0.0026 ⊥	-0.0163	-0.0607	-0.0469	-0.0506
	Unit-H	0.0399 ⊥	0.0095 ⊥	0.0384	0.0103	-0.0058	0.0813
	Unit-I	0.1149	0.1028	0.1377	0.0453	0.1459	0.1359
	Unit-J	0.1382	0.1732	0.2288	0.1312	0.1964	0.1909
	Unit-K	-0.0253 ⊥	0.1043	0.1459	0.0800	0.1475	0.0900
	Unit-L	-0.0437 ⊥	0.0065 ⊥	-0.0090	-0.0647	0.0517	0.0495
	Unit-M	0.0302 ⊥	0.0416	0.0429	-0.0107	0.0553	0.0766
	Unit-N	0.1184	0.1191	0.1132	0.0745	0.0789	0.0785
	Unit-O	0.1126	0.2068	0.2539	0.2090	0.2506	0.2334
Lot size - Region Inter. Dummies	**meter^2-B**	0.1486	0.1874	0.2390	-0.0171	0.2773	0.2327
	meter^2-C	0.2989	0.2440	0.2781	0.2630	0.2447	0.1173
	meter^2-D	-0.1487	-0.1010	-0.0484	-0.1306	-0.0040 ⊥	-0.0186
	meter^2-E	-0.1744	-0.0647 ⊥	-0.0615	-0.1173	-0.0310	-0.0320
	meter^2-F	-0.0680 ⊥	0.2781	-0.0153	0.0055 ⊥	0.1231	0.3294
	meter^2-G	-0.2014	-0.0852 ⊥	-0.1668	-0.1462	-0.0694	-0.1271

Note: ⊥ denotes that the coefficients are not significant at the 5% level.

Appendix table 1: Estimated Coefficients in Six Periods (continued)

	meter^2-H	-0.1214	-0.0948	-0.0674	-0.0760	-0.0710	0.0319
	meter^2-I	0.0427 ⊥	0.0631 ⊥	0.1272	0.0703	0.1653	0.2373
	meter^2-J	-0.1158	-0.0314 ⊥	0.0383	-0.0171	0.0351	0.0522
	meter^2-K	-0.1372	-0.0339 ⊥	0.0133	0.0035 ⊥	0.0581	0.0335
	meter^2-L	-0.1379	0.0650 ⊥	-0.0008	-0.0477	0.0865	0.1279
	meter^2-M	-0.0200 ⊥	0.0597 ⊥	0.0524	0.0453	0.1163	0.1223
	meter^2-N	-0.0528 ⊥	-0.0133 ⊥	-0.0012 ⊥	0.0072 ⊥	0.0266	0.0334
	meter^2-O	-0.1334	-0.0293 ⊥	-0.0081	-0.0093 ⊥	0.0612	0.0745
Dwelling type-	**Unit-1**	-0.0144 ⊥	0.0131 ⊥	-0.0451	-0.0821	-0.0713	-0.0556
Bedroom	**Unit-3**	0.0556	0.0635	0.0773	0.0815	0.0802	0.1086
Inter. Dummies	**Unit-4**	0.1163	0.1604	0.1402	0.1798	0.1801	0.1666
	Unit-5	0.3996	0.4758	0.4496	0.5748	0.6845	0.5054
Lot Size-	**meter^2-1**	0.1657	0.2536	0.2036	0.1161	0.1895	0.0766
Bedroom	**meter^2-3**	-0.0453	-0.0043 ⊥	-0.0243	-0.0478	-0.0295	0.0124
Inter. Dummies	**meter^2-4**	-0.0066 ⊥	0.0095 ⊥	-0.0386	0.0125	0.0165	0.0549
	meter^2-5	0.0749	0.0510 ⊥	0.0013 ⊥	0.0540	0.0910	0.1163
Dwelling type-	**Unit-2**	0.0579	0.0664	0.0722	0.0683	0.0792	0.0936
Bathroom	**Unit-3**	0.2037	0.2576	0.2734	0.1965	0.2512	0.1706
Inter. Dummies	**Unit-4**	0.5703	0.8623	0.7825	0.6470	0.5156	0.4185
Lot Size-	**meter^2-2**	0.0601	0.0616	0.0753	0.0577	0.0479	0.0247
Bathroom	**meter^2-3**	0.1655	0.1635	0.1723	0.1366	0.1127	0.0921
Inter. Dummies	**meter^2-4**	0.1552	0.2270	0.1293	0.1037	0.0891	0.0467
Bedroom-	**3-2**	0.0086 ⊥	0.0158	0.0117	0.0136	-0.0021	0.0107
Bathroom	**3-3**	-0.0364	-0.0361	-0.0726	-0.0759	-0.1132	-0.0323
Inter. Dummies	**4-2**	0.0090 ⊥	0.0288	0.0196	0.0399	0.0191	0.0519
	4-4	-0.0022 ⊥	0.0122 ⊥	0.0055	0.0058	-0.0174	0.0652

Note: ⊥ denotes that the coefficients are not significant at the 5% level.