18TH ANNUAL PACIFIC-RIM REAL ESTATE SOCIETY CONFERENCE

ADELAIDE, AUSTRALIA, 15-18 JANUARY 2012

BUILDING AGE, DEPRECIATION AND REAL OPTION VALUE - AN AUSTRALIAN CASE STUDY

Greg Costello

Greg Costello, Curtin University, Perth, Western Australia

G.Costello@curtin.edu.au

ABSTRACT

This paper examines the influence of building age as a measure of depreciation and real option (redevelopment) value within a large Australian housing market. A number of authors have proposed that building age used as an independent variable in hedonic analysis of housing markets can proxy for other omitted variables. Recent research suggests that depreciation of structures, location influences and real option potential are all measured implicitly with building age if hedonic models are not correctly specified so as to effectively identify and isolate the influence of these factors. The influence of building age in identifying and measuring the option value of redevelopment potential is tested empirically in a large Australian housing market (Perth, Western Australia). Specific housing submarkets are identified as having significant variance in building age of the housing stock. Transactions from these submarkets over the sample period 1995-2010 are analysed within several hedonic models to test these influences. Consistent with recent similar research in this area, an "intensity" variable is used. This variable is constructed from the ratio of building areas from sales of existing properties to building areas from sales of "new" construction within the same submarkets. The preliminary results confirm this variable as providing robust characteristics in identifying and measuring these influences.

****The author gratefully acknowledges Landgate (WA) for providing the data used in this research****

1. Introduction

Since the seminal work of Rosen (1974) hedonic pricing theory has been widely applied to the analysis of housing markets. Typically, hedonic analysis of housing markets has used a general modelling framework represented in matrix and cross-section form:

$$P_i = \beta' X_i + \varepsilon_i \quad (1)$$

Where: P_i represents sale price or "bundle price" of the asset, typically expressed as the natural logarithm of selling price, X_i is a vector of attributes for each transaction, $X_{j,}(j = 1, \dots, n)$, i indexes *M* individual property sales in cross-section each with *n* hedonic characteristics, β so that; $\beta' X_i$ represents an *n*-dimensional column vector of implicit market prices for property characteristics as of time *t*. In a well specified model, ε_i is the disturbance term arising from negotiations between buyers and sellers who have idiosyncratic characteristics (typically assumed to be iid). This general model form can be considered as representing "use" value for property *i* at time *t*.

Housing is a durable asset so when introducing the influence of time to this model several important issues arise, significantly the issue of building depreciation. Building age is an important variable generally included in the vector X_i of attributes for each transaction. Typically, the coefficient on the building age variable is considered as a measure of the influence of depreciation of the structure. A common result for numerous studies in many regions of the world confirms increasing value for older properties (see for example Lee Chung and Kim; 2004, Coulson and McMillen; 2008, Goodman and Thibodeau; 1995). A common explanation for this result is that building age represents a proxy variable for land characteristics. For example, it is common in most cities for older homes to be located in more desirable locations in terms of physical features (ocean or river frontage, elevated sites etc). In addition, these sites will have advantages in the form of access to important infrastructure and may also have larger plot sizes. Hence, the error term, ε_i represented in equation (1) has been frequently discussed as representing omitted influences (variables) of both depreciation in the structure and variations in land characteristics.

More recently, this has come to be discussed as "option value" arising from the fact that the owner of a housing unit has the real option to reconfigure X_i, the vector of attributes for each property through either renovation or demolition and reconstruction on the existing land.

In a significant recent contribution to the hedonic pricing literature for housing market analysis, Clapp and Salavei (2010) advocate the addition of another term to equation (1) for the value of this option, considered as the right but not the obligation to change the characteristic vector X_i . [1]

$$P_i = \beta_1' X_i + \beta_2 I_i + \varepsilon_i \qquad (2)$$

Where: $\beta_1' X_i$ is an *n*-dimensional column vector of implicit market prices for property characteristics as of time *t* representing the "use" value, and $\beta_2 I_i$ is a scalar aggregation index representing "option" value as measured through I, an "intensity" variable specified to measure the amount of structure per unit land value for property *i* at time *t*. When correctly specified, the aggregator variable I should increase with size of the building and quality of other amenities (bedrooms, bathrooms, swimming pools etc) and decrease with building age and with land value. The theory developed by Clapp and Salavei (2010) demonstrates that "use" value should increase with the aggregator variable for new construction where the option value is near zero.[2] Conversely, option value should increase and use value should decrease with older structures where the land is less optimally developed. In the absence of a correctly specified option value term, the age variable will capture option value in that the city-wide hedonic regression will indicate that property value increases with age of the structure.

This paper tests the theoretical framework developed by Clapp and Salavei (2010) with an empirical application (case study) to data from Perth, Western Australia. To my knowledge this is the first application of the methodology used by Clapp and Salavei (2010) to a new source of market data outside the US. This is an important new field of empirical research that needs to be applied in alternate markets as the methodology is dependent upon the development of appropriate "intensity" variable specifications. My results reveal significant consistency with the results of previous research and some variations on the appropriate specification of "intensity" variables. The paper proceeds as follows: in section 2 I provide further detail concerning theoretical framework and model specification, in section 3 I discuss data, in section 4 I review results and in section 5 provide some conclusions and recommendations for further research.

2. Model Specification and Intensity Variables

The theoretical functional form developed by Clapp and Salavei (2010) in equation (2) raises several important issues for empirical research; first, an appropriate method of measurement of intensity, I for the vector, X_i of existing housing attributes and, second the specification of a non-linear regression model required to identify use value and option value.

---

[1] Clapp and Salavei (2010) provide a complete theoretical exposition as motivation for this model. The notation used here represents a simplified version of their theoretical framework.

[2] For example, a property with a newly built structure will typically be near optimum, so the value of the option will be low, whereas a smaller old building on identically valued land will be associated with large option value.

An important identification issue arises with the specification of the intensity (aggregator) variable, I. In the standard hedonic specification for housing market analysis as indicated in equation (1) it is common to include all the characteristics of the structure and the land (lot). The important issue is how to distinguish these characteristics from the construction of the intensity variable, I given that this variable is a positively-valued function of the vector of hedonic characteristics. This is an empirical issue, since the intensity variable will be highly correlated with hedonic characteristics in the sample unless appropriately identified. To overcome this identification issue, I explicitly develop and test two intensity variables as developed by Clapp and Salavei (2010):

Assessed structure value divided by estimated land value for each observation ($Intensity_{Assessed}$). In constructing this variable I use estimated annual land values for individual transactions provided by Landgate, an agency of the government of Western Australia. The procedure involves subtracting the estimated land value from the transaction price to arrive at an estimated structure value for each transacting property. This procedure differs somewhat from that utilised by Clapp and Salavei (2010) who used an explicit assessed structure value from the assessing authority in their data. This variable was not available in my data.

I also use an intensity variable specification which calculates the ratio of building area (interior square metres) of a transacting property to the average building area of nearby new construction ($Intensity_{New\ Build}$). I define new buildings as being six years or younger in building age. In specifying the dimension of "nearby new construction" I impose a distance restriction as specified through a local map grid reference specification. In brief, this results in new construction within a radius of approximately 3 km of a subject property being used in construction of this variable. In addition I impose a six year "time window" around each transaction by including new construction from three years prior a transaction date and also three years after a transaction date. I justify this specification by the fact that market participants contemplating new construction impose their expectations upon market activity and appropriate levels of redevelopment (intensity) of new housing stock. I experimented with a number of specifications for this variable and found that these geographic and time specifications resulted in appropriate sample sizes for this intensity variable specification.

The specification of a suitable regression model to identify and estimate use value and option value introduces a number of important econometric considerations. First, option value will be influenced negatively by structural characteristics whereas these same variables impact positively into the use value portion of the model. As an example, note that building age can enter the model in equation (2) in two ways; (i) in the standard hedonic regression of use value in X_i where age is intended to capture depreciation and, (ii) in the additional intensity variable I_i or the option term, where option value increases with age. This has important implications for cross-sectional hedonic models. The existing level of the hedonic aggregator function I_i varies across different properties, implying different amounts of option value, with the highest option value being in properties with the least amount of I_i.

This characteristic of intensity, I_i introduces special estimation problems associated with the general theoretical functional form indicated in equation (2). The first is an identification issue: one term increases while the other decreases in a desirable characteristic. Second, there are problems associated with measurement of the aggregator variable I_i as distinct from the standard hedonic vector X_i .

In response to these econometric issues, Clapp and Salavei (2010) propose two criteria for appropriate functional form in estimating equation (2). The option value should be additive to the vector of hedonic characteristics and secondly the disturbance term should imply multiplicative errors. Additive option value is an essential feature of this model. It is particularly important where my main intention here is accurate estimation of the implicit market prices of the hedonic (use value) vector. Many properties, especially those newly built will have near zero option value because they are already developed at optimum level for the aggregator variable. In these cases it is important that property value not approach zero and that the standard hedonic use value indicated in equation (1) remains.

A multiplicative error term is assumed when the log of sales price is the dependent variable, a functional form that has become commonplace in the hedonic literature. Semi log or double log hedonic models have the plausible characteristic that errors made by buyers and sellers can be considered as a percentage of underlying value, not constant dollar amounts. Previous empirical work has shown that heteroscedasticity is substantially reduced or eliminated by semi log or double log models. Therefore, I consider only models with log sales price as the dependent variable. As a result of these considerations the general estimation model is of the form:

$$\ln P_i = \alpha_0 + \alpha_1 \ln Built_{area} + \alpha_2 \ln Land_{area} + \alpha_3 Age + \alpha_4 Age^2 + \alpha_5 h + \beta_1 \ln I_i + sB_1^s \ln I_i + \varepsilon_i \tag{3}$$

Where:

$\ln P_i$ is the natural logarithm of selling price for property i .

α_0 is a constant

$\ln Built_{area}$ is the natural logarithm of the building area expressed either in discrete square metres or room count.[3]

$\ln Land_{area}$ is the natural logarithm of the individual land plot area expressed in discrete square metres.

Age and Age^2 represents a polynomial function of building age (Malpezzi et al., 1987).

h is a vector of other characteristics such as car parking facilities, swimming pools, bedroom and bathroom details and time dummies (full description of data in section 3).

---

[3] Our data (discussed in next section) contains a significant number of observations with missing values for building area in square metres. As a result, we use a total room count variable as a measure of building size for the total sample.

$\ln I_i$ Is the natural logarithm of the intensity aggregator variable which is additive to the characteristic vector of housing characteristics and allows the error term to be multiplicative.

$sB_1^s \ln I_i$ is a term which shifts $\beta_1 \ln I_i$ at low levels of the aggregator variable; this is a dummy variable with a value of 1 over the lower 25th percentile of I_i , otherwise zero. This term is included to test the non-linear characteristics of option value.

3. Data

The empirical methodology developed by Clapp and Salavei (2010) suggests that the influence of building age and option value is more likely to be observed empirically within samples of older building structures. Furthermore, this influence should be more pronounced in those areas where there is a significant variation in building ages between individual properties. In this case, the influence of depreciation as measured by building age can be more easily observed. Old properties should exhibit higher depreciation and therefore lower use values and higher option values. At the same time, newer buildings should exhibit lower depreciation, higher use values and lower option values.

In addition to these requirements, in order to adequately test the suitability of various intensity variables, most notably the influence of new building structures ($Intensity_{\ New\ Build}$), it is necessary to identify a sample selection procedure which includes a suitable sample size of new buildings. This Sample selection procedure presented several significant problems best understood by reference to Table 1.

Insert Table here

The sample data was obtained from the Landgate W.A. database with the assistance of the Valuer General Office.[4] The data includes transactions from the period 1995-2011. This sample period was selected as 1995 represents the first year in which accurate estimates of land value are available for all transacting properties in Western Australia. Although I sample through to 2010 the final estimation models only include properties transacting until 2008 due to the "time window" requirements for the new building intensity variable discussed in the previous section.

Table 1 provides a selection of specific suburbs within the Perth Metropolitan region to illustrate sample selection procedure issues. Note from Table 1 that the data is classified according to three subsamples (full sample, building area available, building area not available). This is necessary as a number of transactions did not have the building area variable (measured in discrete square metres) available. In this case, a building room count was used as the proxy variable for building size. In order to minimise the influence of bias within the total sample caused through the omission of this variable, information was collected on the different sample sizes, mean building age and standard deviation of building age within respective subsamples. In addition, the coefficient of variation of building age was also calculated (mean age / Std dev). Finally, a sample stability

[4] Landgate is a statutory authority in the state of Western Australia. Landgate is Western Australia's primary source of land information and geographic data, maintaining the state's official register of land ownership and survey information and being responsible for valuing the state's land and property for statutory purposes.

measure was used in order to identify those suburbs indicating least variability between these key subsample characteristics across the full sample of suburbs for the Perth Metropolitan region. The lowest sample stability statistics indicate suburbs with least variation across specific subsamples. Table 1 displays these subsample characteristics for 13 suburbs across the Perth Metropolitan region with the best samples stability characteristics.

Remembering that building age is a desirable characteristic of the sample selection procedure, the suburb of City Beach has been chosen as a suitable subsample for analysis of the influence of depreciation and option value in this case study. By reference to Table 1 it can be seen that this suburb offers the desirable characteristics of a suitable sample size, and the old average building age (30 years) which is also quite consistent across the subsamples (building area available/not available). In addition, all of the subsamples displayed a suitable coefficient of variation to suggest that this sample is relevant in testing for the influence of building depreciation and option value as outlined in the empirical framework developed by Clapp and Salavei (2010). The sample selection procedure indicated in table 1 also suggests several other suitable suburbs for later analysis. Note the significant variation in building age characteristics across these various suburbs. Also, note the difficulty of sample size whereby some suburbs have significantly lower numbers of total transactions in comparison to others.

The key variables for analysis are defined in Table 2. These variables are used in various estimation models derived from equation (3). More details on data sources, calculation procedures and relevant formulae are contained in Table 2. Relevant summary statistics for the City Beach sample are summarised in Table 3.

Insert Tables 2 & 3here

Table 3 displays key descriptive statistics for some of the variables defined in Table 2. Note that summary statistics are presented for both the full sample of all observations and the smaller subsample of observations where the discrete building area variable is available. Where the same variable is available for both subsamples, results are available for Welch's *t*-test which tests the null hypothesis of no difference between the means of these variables assuming different sample sizes and variance between the two samples. It is evident that the null hypothesis is rejected at levels of statistical significance for the majority of variables tested, confirming different average characteristics of the relevant subsamples. In summary, the full sample displays lower average selling prices, older buildings and older transaction dates than the sample of transactions where building area is available. The only variable where there is no statistically significant difference between means is the land area variable.

The *Built value* $_{Estimated}$ variable requires further discussion. Recall that from section 2, this variable is important for construction of the intensity variables and is an estimate of building value as a component of transaction price (= Price - *Land value* $_{Estimated}$). Note that negative minimum values are displayed in both subsamples for this variable. These negative values have been caused in those transactions where the Landgate estimate of land value is higher than the recorded transaction price.

After an investigation of the causes of negative estimates of built value for some transactions it was discovered that a number of such transactions represented vacant land, non-market or non-arms-length transactions and these were removed from the sample, hence 1456 transactions for observation from the original 1795 indicated in Table 1. However, not all transactions displaying negative estimated building values were removed from the sample. After investigation it was discovered that a number of market transactions of older properties within City Beach displayed the tendency for negative and very low positive estimated building values, confirming significant depreciation of the buildings. In addition, the Landgate estimate of land value occurred only once towards the end of the given year of the transaction, allowing the possibility in years of significant capital appreciation for a property transacting at the start of the year to be influenced by a significant difference in estimated land value at the end of the year from which might have existed at the date of the transaction. These transactions with *Built value* $_{Estimated}$ exhibiting zero or negative values present an important consideration for empirical analysis.

Note from Table 3 these transactions also correspond with negative values for the *Intensity* $_{Assessed}$ and *Ratio* $_{BV\_Price}$ variables. Note also from equation (3) that the general form of estimation model utilises the logarithmic transformation of the intensity aggregator variables. Those observations with zero or intensity values cannot be transformed to a logarithmic form. This necessitates the construction of the Intensity $_{Missing\ BV}$ variable consistent with the procedure adopted by Clapp and Salavei (2010). Those transactions with the *Ratio* $_{BV\_Price}$ in the range -.10 – 0 were considered as missing building values and transformed as the Intensity $_{Missing\ BV}$ dummy variable set. Note that this procedure sets the cut off level for assessment as a missing value at the -10% level according to the *Ratio* $_{BV\_Price}$ variable specification. This procedure is only required for the sample where *Intensity* $_{Assessed}$ is used in estimations. Where the *Intensity* $_{New\ Build}$ form of the aggregator variable is used this procedure is not necessary as negative assessments of intensity do not arise with the assessment procedure used for this variable.

From Table 3 it is also useful to examine the statistical properties of the different estimation forms for the intensity aggregator variables. It can be seen that the mean level of *Intensity* $_{New\ Build}$ is significantly higher than the mean level for *Intensity* $_{Assessed.}$ In addition the standard deviation for *Intensity* $_{New\ Build}$ is significantly lower. This indicates some significant differences in both the scale and variability of the aggregator variables. This fact will have some important implications in the interpretation of results with models using the different specification of the aggregator variable. Finally, this influence is also reflected in the mean levels of the 25^{th} pctl – Ln (*Intensity* $_{Assessed}$) and 25^{th} pctl – Ln (*Intensity* $_{New\ Build}$) variables. Note that the levels at which the 25th percentile of intensity occurs is significantly different between the two samples according to the intensity assessment methods used.

4. Results

Tables 4, 5 and 6 present results for equations estimated with the general form of hedonic model as indicated in equation (3). Table 4 presents results for the full sample using the room count variable (Ln_Rooms) as the measure of building size. Table 5 presents results for the smaller restricted sample where the building area is available measured in discrete square metres (Ln_Built area). As an additional test for the influence of sample selection bias, an additional set of results is reported for the restricted sample using the same model specifications as used for the full sample reported in Table 4. The significance of the different data sets occurs with respect to the different intensity aggregator variables used for model specification. Note that in Table 4 *Ln_Intensity* $_{Assessed}$ is used since there is no accurate building dimension available to specify an alternate intensity variable. In Table 5, *Ln_Intensity* $_{New\ Build}$ is the intensity variable used and in Table 6 *Ln_Intensity* $_{Assessed}$ is used to evaluate the influence of sample selection bias. In all sets of results robust *t* statistics are reported. Standard errors for all estimation models are assumed to be spatially clustered. To control for spatial dependence, standard errors are calculated with Huber-White robust methods using the same cell grid over the City Beach suburb as was used for the construction of the new build intensity variable.

Insert Tables 4,5 & 6 here

In all sets of results, Model 1 provides estimates for the standard hedonic model without terms for the option value as measured by the relevant intensity variable. In all sets of results the standard hedonic model has the expected signs on all explanatory variables. It is evident that the influence of increasing building size increases selling price. The difference in magnitude of this variable for the restricted built area sample (0.417) and the level of statistical significance contrasts with the alternate specification (room count) used for the full sample with a coefficient of (0.131) and a lower level of statistical significance. This provides strong confirmation for the importance of an accurate measure of building size in hedonic models for housing markets. The land area variable also has a significant positive influence on selling price with a more uniform set of results between the different samples.

The $Ratio_{Baths\_Beds}$ variable is designed to identify homes with a superior standard of accommodation and the regression results for Model 1 display a uniform positive influence across all data sets. Similarly, the Cars and Pool variables display uniform positive and significant results across all data sets.

The polynomial specification for the influence of building age confirms a general pattern of house prices declining with building age until approximately 40 years of age and thereafter rising. The difference between magnitudes of the Age coefficient for the full sample compared to the built area sample is likely explained by sample selection bias in that the built area sample includes younger properties as confirmed by the statistical summary in Table 3. Note that in all results for the Age^2 variable the influence is statistically significant and close to zero although positive. This pattern of depreciating building value for approximately 40 years and then rising is consistent with the pattern found in a number of other hedonic studies. Clapp and Salavei (2010) report similar results with depreciation continuing until 45 years in Greenwich Connecticut USA and Coulson and Macmillan (2008) report a similar age function for Chicago, however their analysis revealed a significantly longer period of depreciation (about 70 years).

These results for the age function lend significant support to the theory of real option value. The increasing portion of the age function after approximately 40 years cannot be due to physical and functional depreciation. The upward sloping portion of the age function is likely to be explained by the increasing option to redevelop or renovate these older properties.

In Model 1, the time dummy variables for the years 1997-2008 (omitted year 1996) confirm a general pattern for increasing prices over the 12 year period. This period displays a pattern of significant capital growth confirming an overall increase of approximately 160% in general house price movement over the period. It is evident that there is significant temporal variation in some periods explained by other variables which will be discussed in more detail below.

As discussed in the introduction, intensity by either measure should be inversely related to the amounts of option value revealed in the hedonic regression. Model 2 in all tables introduces the intensity variable designed to test for the influence of option value. In Table 4, these variables are *Ln_Intensity* $_{Assessed}$ and Intensity $_{Missing\ BV}$. The expected sign of the *Ln_Intensity* $_{Assessed}$ coefficient is negative with the Intensity $_{Missing\ BV}$ dummy variable expected to provide a positive and significant coefficient. As can be seen in Table 4, these results are contrary to a priori expectations. The intensity variable exhibits a statistically significant positive coefficient. For the corresponding regression with the restricted built area sample in Table 5 the coefficient for the *Ln_Intensity* $_{New\ Build}$ variable is as expected, being negative and statistically significant. In Table 6, the results are consistent with those in Table 4, confirming that this discrepancy is not likely to be the result of sample selection bias. These differences in results provided by different specification of the intensity variables are most likely caused by significant variations in the scale of these variables as indicated in the statistical summary from Table 3. Intuition also suggests that intensity is more likely to be closely correlated with building dimension relative to land area. Interestingly, the *Intensity* $_{New\ Build}$ variable is much more highly correlated with both Built Area (0.958) and Rooms (0.665) than *Intensity* $_{Assessed}$ (0.507) and (0.415).

The third regression (Model 3) is designed to test whether the influence of intensity is nonlinear, with the expectation that most option value is concentrated within the lower levels of intensity. In all model specifications the nonlinear effect of intensity is confirmed with the significant negative coefficient on the log of intensity up to its 25th percentile point. Again there are significant differences between the results for the full sample and restricted sample data sets arising from the specification of intensity variables.

Model 4 is the same specification as the third regression except that the log of intensity is now interacted with the age variable in order to evaluate the influence of option value effects independent from depreciation influences. The results for this regression show that the log of intensity significantly changes the influence of the age effect although this influence is only statistically significant for the restricted built area sample shown in Table 5. Note that when this interaction variable is included the negative of influence (depreciation) from the age variable increases in the influence of intensity becomes insignificant. As argued by Clapp and Salavei (2010) these changes in the age effect are consistent with option value theory. The fact that these influences are only statistically significant when the *Intensity* $_{New\ Build}$ variable is used suggests that with this data sample this is a superior specification of the intensity variable in terms of revealing option influences.

Model five interacts large lot size (a dummy variable identifying plots with land area in the 90th percentile) with the intensity variables. These regressions are intended to test the proposition that large lots are associated with a significant increase in the effect of intensity. The lack of statistical significance in coefficients for these variables suggests little influence of lot size interacting with intensity.

An interesting feature of these results is the significant variation in time dummy coefficients from models where intensity variables are included. For example by reference to Table 4 it can be seen that the coefficient for 1998 varies by approximately 3% between model 1 and model 2 when intensity variables are introduced. Note that significant variation is also evident in other years, notably 2002 and 2008. Importantly these influences are consistent irrespective of which measure of intensity has been used and across all data sets. In addition, the R2 for all models confirm significant explanatory power yet there is still significant variation in the time coefficients between different model specifications. This result confirms that the inclusion of intensity variables is warranted in estimating hedonic price indexes for small local markets. The assumption of "constant quality" would appear to be violated unless temporal periods for the index exhibit similar characteristics in terms of both the depreciation in building stock and likely variations in option value arising from these influences. The results in this study indicate that in local areas where significant variation in age of the building stock and potential for renovation exist there will be periods in the cycle where these influences are significant and may bias price indexes estimated with standard hedonic price methods not including appropriate intensity variable specifications.

5. Conclusions

This paper examines the influence of building age as a measure of depreciation and real option (redevelopment) value within a suburb (City Beach) for a large Australian housing market (Perth, Western Australia). The approach used closely follows some aspects of the methodology adopted by Clapp & Salavei (2010). Consistent with a number of previous authors, the results confirm that building age used as an independent variable in hedonic analysis of housing markets can proxy for other omitted variables, notably the influence of building age and associated real option characteristics.. The results in this study suggest that depreciation of structures, location influences and real option potential are all measured implicitly with building age if hedonic models are not correctly specified so as to effectively identify and isolate the influence of these factors. Importantly, the time coefficients for hedonic price indexes for local housing markets can be influenced significantly by these influences. The most significant measure of intensity (aggregator variable) used to identify the influence of real option value is a variation of the method used by Clapp and Salavei (2010) whereby intensity is assessed according to a ratio of existing building to the average size of new buildings within close proximity of a transacting property. These results are consistent with the theoretical framework suggesting increasing real option value associated with redevelopment potential of the land for older residential properties. Finally, these results have been obtained with a limited empirical study, however results suggest significant potential for further testing of these models within this rich data set for housing market transactions in Western Australia.

References

Clapp, J.M., Salavei, K., 2010. Hedonic pricing with redevelopment options: A new approach to estimating depreciation effects. Journal of Urban Economics 67, 362-377.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy 82, 34–55.

Goodman, A.C., Thibodeau, T.G., 1995. Age-related heteroskedasticity in hedonic house price equations. Journal of Housing Research 6, 25–42.

Coulson, E., McMillen, D., 2008. Simultaneous estimation of cohort, age, and time effects in housing prices. Journal of Housing Economics 17, 138–151.

Lee, B., Chung, B.C., Kim, Y., 2004. Building age, redevelopment, and housing prices: the case of apartment complexes in Seoul. The Journal of Real Estate Finance and Economics 30, 55-80.

Malpezzi, S., Ozanne, L., Thibodeau, T.G., 1987. Microeconomic estimates of housing depreciation. Land Economics 63, 372–385.

Table 1: Sample Selection Process

	Full sample				Building Area Available				Building Area Not Available				Sample Stability
Suburb	N	Mean Age	S.D.	C.V	N	Mean Age	S.D.	C.V.	N	Mean Age	S.D.	C.V.	
STRATTON	1979	9.82	4.16	.42	1238	9.41	4.00	.43	741	10.51	4.33	.41	.20
CITY BEACH	1795	30.06	12.79	.43	1053	26.76	12.05	.45	742	34.74	12.35	.36	.21
ROLEYSTONE	1190	23.13	11.47	.50	469	25.03	11.77	.47	721	21.90	11.11	.51	.24
ERSKINE	1044	7.64	4.35	.57	576	6.58	3.81	.58	468	8.94	4.61	.52	.25
NORANDA	1862	19.69	6.80	.35	737	17.35	6.97	.40	1125	21.23	6.23	.29	.29
ATWELL	2219	6.71	4.31	.64	1463	5.74	4.10	.71	756	8.60	4.07	.47	.33
ILUKA	1201	6.39	3.70	.58	836	5.85	3.36	.57	365	7.63	4.12	.54	.36
KINROSS	2648	7.82	4.14	.53	1118	6.24	3.32	.53	1530	8.98	4.30	.48	.37
CURRAMBINE	2575	7.91	4.10	.52	1007	6.47	3.40	.53	1568	8.84	4.24	.48	.39
MEADOW SPRINGS	1574	7.11	4.85	.68	869	5.12	3.93	.77	705	9.56	4.75	.50	.42
LANDSDALE	1376	6.23	3.82	.61	940	5.39	3.32	.62	436	8.04	4.19	.52	.43
BALLAJURA	7052	12.06	5.89	.49	2701	9.16	4.71	.51	4351	13.86	5.83	.42	.45
HALLS HEAD	4670	12.89	9.79	.76	2078	8.95	7.67	.86	2592	16.05	10.15	.63	.46

Table 2: Variable Definitions

Variable Name	Variable Description	Data Source
25th pctl – Ln (*Intensity* $_{Assessed}$)	Natural logarithm of *Intensity* $_{Assessed}$ if in the lower 25th percentile; otherwise 0	Calculated
25th pctl – Ln (*Intensity* $_{New\ Build}$)	Natural logarithm of *Intensity* $_{New\ Build}$ if in the lower 25th percentile; otherwise 0	Calculated
Age	Age of the property in years (= sale year - year built)	Calculated
Age^2	Age squared	Calculated
Built area	Building area of the structure in square metres	Landgate
Built value $_{Estimated}$	Estimate of building value as component of transaction price (= Price - *Land value* $_{Estimated}$)	Calculated
Cars	Number of car parking bays for property	Landgate
Intensity $_{Assessed}$	Estimated value of building divided by estimated value of land	Calculated
Intensity $_{Missing\ BV}$	Equals 10 when Ln (*Intensity* $_{Assessed}$) is set to -10 because assessed value of the building ~ 0 (Discuss 10 pc requirement)	
Intensity $_{New\ Build}$	The ratio of Built area (sqm) of subject property to the average Built area (sqm) of new construction (<= 6 years) located within ~ 3 km of the subject property and sold within 3 years of sale of the subject property (six year time window; -3 years t-sale +3 years)	Calculated
Land area	Land area of the subject lot in square metres	Landgate
Land value $_{Estimated}$	Estimate of land value for year of transaction	Landgate
Ln(...)	Natural logarithm	
Lot Large Dum	Equals one if lot size is within the highest 10th percentile (>90 pctl) otherwise zero	Calculated
Pool	Dummy variable representing swimming pool existing at transacting property	Landgate
Price	Price at which the property was sold	Landgate
Ratio $_{Baths\_Beds}$	Ratio of bathrooms to bedrooms (= baths / beds) for each transacting property	Calculated
Ratio $_{BV\_Price}$	Ratio of estimate of building value to sale price (= *Built value* $_{Estimated}$ / Price)	Calculated
Rooms	Total room count for each transacting property	Calculated
Sale year	Year of transaction	Landgate
Site ratio	Ratio of built area to land area (= Built area / Land area)	Calculated
Y_1997 ... Y_2008	Time dummy variables representing year of sale, omitted base year 1996	Calculated
Year built	Year of construction	Landgate

Table 3: Summary Statistics – City Beach

Variable	Full sample					Built Area Available					t tests	
	N	Min	Max	Mean	S.D.	N	Min	Max	Mean	S.D.	t	prob
Price	1456	236,000	4,900,000	912,158	623,613	873	252,500	4,900,000	1,018,411	686,043	3.74	0.00
Land area	1456	452	2,572	899	145	873	452	2,572	906	154	0.99	0.32
Year built	1456	1922	2008	1972	12	873	1925	2008	1975	12	6.48	0.00
Age	1456	1	83	30.5	13	873	1	83	27.5	12	5.82	0.00
Sale year	1456	1996	2008	2001	3	873	1996	2008	2002	4	2.05	0.04
$Land\ value_{Estimated}$	1456	209,000	2,805,000	666,507	399,544	873	211,750	2,673,000	711,624	413,001	2.58	0.01
$Built\ value_{Estimated}$	1456	-125,000	3,415,000	245,651	316,069	873	-79,000	3,415,000	306,787	369,497	4.08	0.00
Rooms	1456	5	22	9.7	3	873	5	22	10.6	2	7.93	0.00
Cars	1456	0	5	1.8	1	873	0	5	1.9	1	5.18	0.00
$Intensity_{Assessed}$	1456	-0.09	3.03	0.36	0.32	873	-0.09	3.03	0.42	0.35	4.01	0.00
Built area	873	93	1,129	273	97	873	93	1,129	273	97		
$Ratio_{BV\_Price}$	1456	-0.10	0.75	0.23	0.14	873	-0.10	0.75	0.26	0.15	4.43	0.00
Site ratio	873	0.11	0.80	0.31	0.11	873	0.11	0.80	0.31	0.11		
$Ratio_{Baths\_Beds}$	1456	0.17	2.00	0.46	0.19	873	0.17	1.33	0.48	0.20	3.06	0.00
$Intensity_{New\ Build}$	873	0.22	3.35	0.76	0.27	873	0.22	3.35	0.76	0.27		
25^{th} pctl – Ln ($Intensity_{Assessed}$)	1456	-10.00	0.00	-0.94	2.11	873	-10.00	0.00	-0.68	1.87	3.06	0.00
25^{th} pctl – Ln ($Intensity_{New\ Build}$)	873	-1.52	0.00	-0.19	0.33	873	-1.52	0.00	-0.19	0.33		

Table 4: Hedonic Model Results – Full Sample

Variable	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*
(Constant)	11.358	39.0	12.349	48.6	12.522	51.0	12.570	50.9	13.056	42.6
Ln_Rooms	0.131	4.1	0.070	2.5	0.070	2.6	0.080	2.9	0.083	3.0
Ln_Land area	0.213	5.1	0.099	2.7	0.081	2.3	0.073	2.1	-0.002	0.0
Age	-0.025	-15.8	-0.016	-11.3	-0.013	-9.0	-0.014	-8.5	-0.014	-8.4
Age^2	3.20E-04	13.9	2.14E-04	10.5	1.74E-04	8.7	1.82E-04	7.5	1.76E-04	7.3
Cars	0.053	5.9	0.040	5.1	0.034	4.6	0.035	4.7	0.033	4.5
Pool	0.072	5.3	0.044	3.7	0.042	3.8	0.043	3.8	0.044	4.0
$Ratio_{Baths\_Beds}$	0.173	4.8	0.134	4.3	0.119	4.0	0.117	3.9	0.120	4.0
Y_1997	0.092	3.0	0.099	3.8	0.091	3.6	0.092	3.6	0.093	3.6
Y_1998	0.208	6.7	0.170	6.4	0.165	6.4	0.165	6.5	0.166	6.5
Y_1999	0.350	11.8	0.327	12.8	0.315	12.7	0.314	12.7	0.316	12.8
Y_2000	0.445	13.9	0.428	15.6	0.427	16.1	0.426	16.1	0.428	16.2
Y_2001	0.538	17.2	0.547	20.3	0.549	21.1	0.545	20.9	0.550	21.0
Y_2002	0.667	21.7	0.717	27.0	0.715	27.9	0.713	27.8	0.715	27.9
Y_2003	0.802	26.0	0.792	29.9	0.790	30.9	0.786	30.6	0.788	30.8
Y_2004	0.938	29.4	0.954	34.8	0.931	35.1	0.927	34.8	0.928	34.9
Y_2005	1.105	34.1	1.102	39.5	1.094	40.7	1.089	40.3	1.092	40.4
Y_2006	1.399	41.4	1.414	48.5	1.396	49.6	1.390	49.1	1.390	49.1
Y_2007	1.625	45.8	1.638	53.6	1.638	55.6	1.634	55.3	1.635	55.4
Y_2008	1.647	40.4	1.589	45.2	1.565	46.0	1.560	45.8	1.565	46.0
$Ln\_Intensity_{Assessed}$			0.137	21.3	0.239	20.7	0.268	13.7	0.269	13.7
$Intensity_{Missing\ BV}$			0.094	15.3	0.096	16.3	0.094	15.6	0.093	15.5
25^{th} pctl – Ln ($Intensity_{Assessed}$)					-0.090	-10.5	-0.085	-9.6	-0.084	-9.4
Age * $Ln\_Intensity_{Assessed}$							-0.002	-1.8	-0.002	-2.0
Age^2 * $Ln\_Intensity_{Assessed}$							0.000	1.5	0.000	1.7
Lot Large Dum * $Ln\_Intensity_{Assessed}$									0.013	1.3
Lot Large Dum									0.077	2.9
N	1456		1456		1456		1456		1456	
Adj R2	0.834		0.877		0.886		0.886		0.886	

Table 5: Hedonic Model Results – Built Area Sample

Variable	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*
(Constant)	9.037	14.2	6.107	7.9	6.329	8.1	6.645	7.9	7.118	9.1
Ln_Built area	0.417	9.5	0.921	6.7	0.917	6.6	0.954	6.5	0.958	6.5
Ln_Land area	0.264	2.8	0.259	2.9	0.229	2.6	0.162	1.8	0.088	0.6
Age	-0.016	-4.4	-0.017	-4.9	-0.016	-4.6	-0.022	-5.1	-0.022	-5.2
Age^2	1.94E-04	3.6	1.99E-04	3.7	1.84E-04	3.5	2.81E-04	3.0	2.77E-04	3.0
Cars	0.025	1.7	0.024	1.7	0.024	1.8	0.027	2.2	0.027	2.2
Pool	0.047	2.7	0.048	2.9	0.047	3.0	0.048	3.6	0.050	3.9
$Ratio_{Baths\_Beds}$	0.101	2.3	0.092	2.1	0.086	2.0	0.068	1.5	0.071	1.5
Y_1997	0.067	1.9	0.063	1.7	0.062	1.7	0.068	1.9	0.066	1.9
Y_1998	0.181	5.5	0.169	5.2	0.175	5.4	0.164	5.0	0.159	4.8
Y_1999	0.278	8.8	0.236	6.6	0.234	6.7	0.224	6.4	0.225	6.6
Y_2000	0.401	11.2	0.406	11.0	0.408	11.5	0.402	10.5	0.400	10.5
Y_2001	0.469	16.4	0.466	14.2	0.475	15.9	0.459	13.5	0.461	14.0
Y_2002	0.584	20.8	0.641	25.7	0.639	26.8	0.622	23.6	0.622	23.8
Y_2003	0.737	23.9	0.797	25.5	0.800	27.6	0.776	23.2	0.774	24.2
Y_2004	0.869	31.1	0.917	36.9	0.921	37.8	0.899	31.2	0.898	31.3
Y_2005	1.014	31.9	1.056	35.3	1.052	35.9	1.033	31.3	1.033	31.3
Y_2006	1.339	27.4	1.363	30.2	1.369	32.8	1.345	32.0	1.343	31.2
Y_2007	1.534	45.2	1.542	36.9	1.547	41.3	1.530	40.4	1.528	40.2
Y_2008	1.555	31.8	1.565	31.2	1.567	32.1	1.554	32.4	1.552	31.8
$Ln\_Intensity_{New\ Build}$			-0.533	-3.7	-0.434	-3.0	0.099	0.4	0.155	0.6
25^{th} pctl – Ln ($Intensity_{New\ Build}$)					-0.117	-2.4	-0.042	-1.1	-0.040	-1.2
Age * $Ln\_Intensity_{New\ Build}$							-0.035	-3.4	-0.037	-3.7
Age^2 * $Ln\_Intensity_{New\ Build}$							0.000	2.5	0.000	2.7
Lot Large Dum * $Ln\_Intensity_{New\ Build}$									-0.101	-1.1
Lot Large Dum									0.028	0.5
N	873		873		873		873		873	
Adj R2	0.866		0.870		0.871		0.875		0.875	

Table 6: Hedonic Model Results – Restricted Sample

Variable	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*	Coefficient	Robust *T*
(Constant)	10.312	15.2	11.505	15.7	11.742	16.2	11.722	16.2	11.871	11.5
Ln_Rooms	0.170	3.2	0.112	3.4	0.109	3.0	0.106	3.0	0.107	3.0
Ln_Land area	0.361	3.3	0.220	1.9	0.193	1.7	0.196	1.7	0.173	1.1
Age	-0.021	-6.1	-0.013	-6.9	-0.010	-5.8	-0.008	-4.2	-0.008	-3.9
Age^2	2.11E-04	4.92	1.24E-04	5.5	8.20E-05	3.4	5.62E-05	1.7	5.53E-05	1.6
Cars	0.052	2.6	0.034	1.9	0.030	1.8	0.030	1.8	0.029	1.7
Pool	0.078	6.1	0.057	7.7	0.053	6.7	0.052	6.3	0.052	6.3
$Ratio_{Baths\_Beds}$	0.161	3.4	0.116	3.3	0.104	3.0	0.105	3.1	0.106	3.2
Y_1997	0.060	1.7	0.070	3.1	0.056	2.4	0.056	2.4	0.055	2.3
Y_1998	0.194	5.8	0.162	5.6	0.148	5.5	0.148	5.7	0.147	5.7
Y_1999	0.301	9.6	0.271	10.9	0.252	9.2	0.252	9.1	0.251	9.2
Y_2000	0.418	10.2	0.401	10.6	0.396	10.3	0.397	10.4	0.397	10.6
Y_2001	0.483	14.1	0.492	14.8	0.488	14.0	0.489	13.9	0.489	13.9
Y_2002	0.608	16.2	0.678	24.6	0.667	25.9	0.668	26.4	0.667	25.5
Y_2003	0.760	19.5	0.742	24.6	0.728	25.6	0.729	25.2	0.728	25.8
Y_2004	0.899	27.0	0.911	31.2	0.882	29.1	0.882	29.1	0.881	28.9
Y_2005	1.047	26.3	1.037	26.4	1.025	26.4	1.026	26.5	1.026	26.9
Y_2006	1.362	23.2	1.383	29.3	1.358	31.6	1.360	31.7	1.358	31.3
Y_2007	1.574	50.6	1.581	54.5	1.584	47.1	1.582	46.1	1.582	45.8
Y_2008	1.627	37.3	1.567	39.5	1.543	41.1	1.547	40.7	1.547	41.0
$Ln\_Intensity_{Assessed}$			0.160	9.8	0.243	13.5	0.244	7.9	0.243	7.4
$Intensity_{Missing\ BV}$			0.113	7.9	0.106	8.2	0.106	8.7	0.105	7.6
25^{th} pctl – Ln ($Intensity_{Assessed}$)					-0.084	-7.7	-0.084	-7.3	-0.083	-7.4
Age * $Ln\_Intensity_{Assessed}$							0.000	0.2	0.000	0.2
Age^2 * $Ln\_Intensity_{Assessed}$							0.000	-0.8	0.000	-0.8
Lot Large Dum * $Ln\_Intensity_{Assessed}$									0.011	0.3
Lot Large Dum									0.027	0.4
N	873		873		873		873		873	
Adj R2	0.839		0.891		0.898		0.898		0.898	