THE REVISITED OF MALAYSIAN HOUSE PRICE INDEX

Norshazwani Afiqah* Mohd Lizam Abdul Jalil*****

Faculty of Technology Management, Business and Entrepreneurship
Universiti Tun Hussein Onn Malaysia
86400 Parit Raja Batu Pahat Johor
Malaysia
hp110062@siswa.uthm.edu.my* lizam@uthm.edu.my jalil@uthm.edu.my*****

ABSTRACT

Price of residential properties plays a very important role in the economics of developed countries. It is an essential factor for individual or household to decide on selling and buying properties and to invest in the direct property market. Of these reason, the Valuation and Property Services Department (VPSD) has constructed the Malaysian House Price Index (MHPI). However, there is no other house price index that could be used as a comparison to the existing MHPI. Thus, parties that have interest in real estate market could not compare the house price index. In this study, an alternative house price index called as the hypothetical price index is constructed. Method used to construct the hypothetical price index is the hedonic method. This method will be performed by using regression technique in which it comprises both of the dependent and independent variables. The hypothetical price index is constructed by using the time-variant approach. Variables used in the model are properly selected and regression with respect to both functional form and independent variables should also be defined correctly in order to obtain unbiased estimates for the house price index. The hypothetical price index shows a same trend as the MHPI. Only at certain quarters, the index is different. It might be due to the different sample of data. It is hoped that this study could contribute in producing an alternative house price index to be used as one of the reference in monitoring house price changes.

Keywords: House price index, hedonic method

Introduction

The advancement of technology provides the opportunities for one to understand the dynamic of real estate market. Similarly, it also enables continuous effort in producing an accurate measure of real estate price indicator. One of the real estate price indicators that is widely use is the house price index. The purpose of house price index is to provide an overview on the price changes over time. The presence of house price index is very important as Wallace (1996) noted that the residential market has its own cycle where it is prone to boom and bust. In other word, the cycle of housing market can be said as unpredictable and it lead to the volatility of real estate market.

Due to this, the house price index has become an important tool to parties that involve in real estate market. Parties that use the house price index extensively are investors, financial institution, researchers and developers. According to Longford (2009), house price index plays an important role for individual to decide on buying or selling property. Besides, financial institution also certainly used the house price index to monitor the property price stability as part of the lending are backed by property.

Most of developed countries such as United Kingdom (UK) and United States (US) have produced their own real estate price index. Lim and Pavlou (2007) stated that the application of the house price index in UK could be seen as early as 1973. The house price index was first produced by the mortgage providers and recently it has been produced by the government sector, the Land Registry. The house price index produced by the government sector is known as the Land Registry House Price Index, whilst the one that is produced by the private sector are the Halifax House Price Index and Nationwide House Price Index.

US also produced house price index due to the needs in monitoring the real estate price changes. For instance, US Federal Housing Finance Agency introduced house price index to measure the movement of house prices for single family. Besides, Freddie Mac House Price Index (FMHPI) was introduced since 1975 in order to measure the house price inflation.

As for Malaysia, the effort to produce house price index began in 1993. The house price index is known as the Malaysian House Price Index (MHPI) is a product of the Valuation and Property Services Department (VPSD). Norhaya et al., (2008) noted that the purpose of MHPI is to monitor the changes of real estate prices from a period to another. It is constructed in order to help in designing the economic national policy for the property development.

The construction of MHPI is by adopting the hedonic method. Basically, the hedonic method is a widely used method for constructing house price index. In US, hedonic method has been used as early as 1968 by US Census Bureau to construct a house price index. On the other hand, in UK the method is applied to construct the Nationwide and Halifax house price index since 1983.

The hedonic method is one of the methods used to construct real estate price index. This method is applicable for a transaction based data. It has been introduced by Griliches and in 1974 it was formalised by Rosen. According to Rosen (1974), based on the hedonic hypothesis, products or goods are valued based on their characteristics. For instance, price for a house will be valued according to its characteristics or attributes such as number of bedrooms, building area, land area and others. The hedonic method can be used for different types of product which the value for each of the product will be based on the characteristics of the product. This is due to the heterogeneity of the product.

It could also be seen from the definition of the hedonic prices by Rosen (1974) which he defined the hedonic prices as an implicit price of its attributes that being determined by the heterogeneity and the characteristics of the observed products. The implicit price is the estimated price obtained from the first-step regression analysis in the construction of hedonic price indices. The regression analysis involved two types of variables which are the dependent and the independent variables. The dependent variable is the price of the products or goods and the independent variable is the characteristics of the goods.

There are some advantages that can be obtained by applying the hedonic method in constructing the house price index. One of it is the house price index can be easily constructed as this method is relying on larger amount of transaction data. Other than that, the method has the ability to embed two important attributes of the property in estimating the property's value (Tan, 2011). The two important attributes are the physical and locational characteristics of property. For instance, to determine a value of a house, ones can consider the physical characteristic of the house and the location characteristics within the same regression.

The application of regression analysis in the method helped one to determine the attributes that give a high impact on the property values. Thus, one can differentiate which attribute is actually contributed to the house prices. The example of hedonic model given by Ramanathan (2002) is as shown below. The model shows that the attributes of the properties has been unbundled when applying the hedonic method.

$$Price_i = \beta_i + \beta_2 LandA + \beta_3 BuildA + \beta_4 Bedr + \beta_5\ Bath + \varepsilon_1 \qquad \text{(eq. 1)}$$

Where;
$Price_i$ = Price of property
β_i, β_2,... β_5 = Coefficient parameter that will be determined
LandA = Area of land for the property
BuildA = Area of building for the property
Bedr = Number of bedrooms
Bath = Number of bathrooms
ε_1 = Residual error

In order to use the hedonic method in constructing the house estate price index, it is important to carefully determine the variables that contribute to the price of the properties. Subsequent effect, if the essential variables are not being identified, it will lead to omitted variables bias.

Methodology

Data

This study is based on the transaction data of residential property. The data is originated from the National Property Information Centre (NAPIC) containing residential property transaction of double-storey terrace houses. A total of 5,365 transaction covering Kuala Lumpur and its surrounding areas were available for this study. However, during the data cleaning process it was found that there are inconsistencies and missing values in the data. As a result, only 3,980 transaction data are available. Observations with missing values are removed for the construction of house price index.

Method to construct the house price index

Method used to construct the house price index is the hedonic method. Two approaches are available in adopting the hedonic method. The approaches are the exact hedonic equation and the time-variant index approach. In this study, the time-variant index approach is applied due to the availability of data.

In order to adopt hedonic method, multiple regression technique is used. Two types of variables are needed to apply this technique. These are the dependent and the independent variables. Transaction price of house represents the dependent variable whilst the house characteristics represent the independent variable.

The independent variables used in this research could be divided into three categories. These are the physical characteristics of property, locational characteristics and time-dummies. First category of independent variable; the physical characteristics of property, it usually will describe the building structure. Past studies of Case and Szymanoski (1995), Osland (2010) and Dorsey et al., (2010) have shows that there are no specific variables when modelling house price. Nevertheless, variables that mainly used to describe the physical characteristics are as listed below:

Table 1: Physical characteristic of property

Physical characteristics	Description
Size	Size of house could be divided into two categories. These are the lot area and built up area. Lot area represent the size of land on which the house resides, whilst built up area represent the overall building size.
Age	Age of building is used to represent the depreciation that occurred to the building. It is measured by taking the difference between the building transaction year and its completion year.
Bedroom	The variable represents the number of bedroom that is available in the subject property. It is expected that the increase of bedrooms' number will increase the property prices.

The second category of independent variables is the locational characteristics. It is also one of the important attributes that contribute to the house price. This is supported by Kiel and Zabel (2008), characteristics of house that affect house prices are spatially related in the form of "location, location, location" hierarchy. Studies done by Brunauer (2010) and Lehner (2011) have used market segmentation and property linkage to describe the locational characteristics of property respectively. In this study, market segmentation and property linkage are adapted to describe the locational characteristic of property. Details of the location characteristics are listed below:

Table 2: Locational characteristics of property

Locational characteristics	Description
Sub-districts	There are a total of five sub-districts used in this study. These are sub-districts Batu, Cheras, Kuala Lumpur, Petaling and Setapak.
Property linkage	The property linkage is described as the proximity or accessibility to specific externalities. This study used distance to the nearest city centre to represent the property linkage. It is measured in kilometres.

As has been stated earlier, this study is based on the second approach of hedonic method. That is the time-variant index approach. Thus, time-dummies also become one of the independent variable used in this study. The construction of house price index involved transaction data from year 2005 to 2012. Quarterly time-dummies are used as variable. Total quarters are 30 which it started from quarter 1, year 2005 until quarter 2, year 2012.

Thus, the relationship between the two variables; dependent and independent variables are as shown in the equation below:

Price = f (physical characteristics, locational characteristics, time dummies) (eq. 2)

The equation of multiple regression analysis for the above model is:

$$Y_i = \alpha + \sum_{k=1}^{K} \beta_{ki} + \beta_2 X_{i2} + \beta_3 X_{i3} + \ldots + \beta_t X_{it} + \varepsilon_i \qquad \text{(eq. 3)}$$

Where i = 1, 2, …, n (n = number of observations), Y_i = house price for ith transaction, $\beta_1,\ldots,\beta_i$ = determined coefficient parameter, X_{it} , t = 2,…, n. β_{ki} , k represent the previous mentioned property characteristics which k= 1, 2,…k. t is the time dummies variable where a dichotomous value is used to represent when transaction of house occur in that quarter. Value of 1 is used to represent the existing of the transaction in a quarter whilst 0 values represent that there is no transaction of property in that quarter. The price index in period t is given by computing the value of anti-logβ_t.

Equation 3 is used to model the house price index by specifying the details of attributes used. The model used to construct the house price index is given below:

$$\log Price_i = \alpha + \sum_{k=1}^{K=5} \beta_{ki}\ subdistrict\ + \beta_2 \log LotArea + \beta_3 \log BuiltArea +$$
$$\beta_4 \log Bedroom + \beta_5 \log Distance + \beta_6\ Age + \sum_{k=1}^{K=30} \beta_{ki}\ Time + \varepsilon_i \qquad \text{(eq. 4)}$$

Where k represents the number of parameters in each variable, i represents the number of observation (i = 1, 2,…, n). α represent the constant value of a model, ε is the error terms and β_i (i = 1, 2,…, n) represent the coefficient for each parameter used.

Equation 4 represents the hedonic equation in the construction of house price index. To ensure the model is robust, the ordinary least square (OLS) technique is employed in order to produce an unbiased estimated for each of the variable's coefficient.

The OLS assumptions that need to be followed in modeling the house price index are listed below:

i. The dependent variable are linear function of X plus random error term (residual).
ii. Independent variable and error term (residual) are uncorrelated.
iii. Error term produce zero mean and constant variance (homoskedasticity).
iv. Error terms (residuals) are serially uncorrelated.

Result and Discussions

Descriptive statistics of dependent and independent variables

This section will discuss the result of the model constructed. The descriptive statistics of the dependent and independent variables used in this study are as shown below:

Table 3: Descriptive statistics of the independent variables

Variables	Mean	Standard deviation	Minimum	Maximum
Dependent variable				
-Transaction price (RM)	624,222.50	278,273.80	270,000.00	1,620,000.00
Independent variables				
-Lot area (sq. m)	183.23	58.65	130.00	647.14
-Built up area (sq. m)	179.29	31.72	148.00	449.56
-Building age (year)	19.65	12.08	1	52
-Number of bedrooms	3.62	0.66	3	6
-Distance to city centre (km)	12.70	3.33	3	23

As shown in the table 3 above, the minimum of house prices used in this study is RM 270,000 and the maximum house price is RM 1,620,000. For the lowest house price, it represents the house with a small land area and built-up area, whilst highest price of house normally represent the house that comes with larger land area and built-up area. The mean for the house price is RM 624,222. 50.

For lot area, the minimum size is 130.00 sq. m and the maximum size is 647.14 sq. m. For the minimum size, it normally represents double-storey terraced house located in the intermediate lot whilst for the larger size it normally represents the corner lot houses. Same goes to the built up area, which the size normally is depending on the location of house whether it is intermediate, corner or lot house. The minimum of the built-up area is 148.00 sq. m and the maximum is 449.56 sq. m.

For building age, the maximum age is 52 years whilst the minimum is 1 year. The mean which is 19.65 years indicates that most of the houses included in the data will have a significant effect on the house prices. There will be depreciation in the house value when the building age is increasing. On the other hand, for number of bedrooms, it normally depends on the size of the house. Thus, the maximum number of bedroom, 6 indicates the size of house is bigger than the minimum number of bedroom, three.

Next variable of property linkage, the minimum distance from house to the city centre is 3 km and the maximum distance is 23 km. The mean for the distance of subject house to city centre is 12.70 km. This indicates that most of the houses are located nearer to its city centre. Below show the other variable that describes the locational characteristic of property; the sub-districts:

Table 4: List of Sub-districts

District	Sub-districts	Number of transaction	Percentage
Kuala Lumpur	Batu	1311	32.93
	Cheras	31	0.78
	Kuala Lumpur	1574	39.55
	Petaling	926	23.27
	Setapak	138	3.47

Table 4 above shows the highest number of transaction comes from sub-district Kuala Lumpur with the percentage of 39.55 out of 100 whilst the lowest number of transaction comes from sub-district Cheras with only 0.78 percent. The highest number of transaction in Kuala Lumpur area is because, it is one of the mature sub-districts and it is most developed area. As a result, many development of house takes place in that area and it contributes to the highest transaction of property.

Besides, time dummies are also used as one of the independent variables to construct the house price index. In this study, total quarters used are 30 which it started from Quarter 1, year 2005 until Quarter 2, year 2012. Table 5 below show the total numbers of transaction data for each quarter involved:

Table 5: Number of transaction data based on quarter

Year / Quarter	Q1	Q2	Q3	Q4	Total
2005	108	126	127	97	458
2006	106	121	142	143	512
2007	115	154	161	61	491
2008	90	104	84	97	375
2009	102	194	240	310	846
2010	114	198	157	175	644
2011	130	156	105	104	495
2012	93	66	-	-	159
TOTAL					3980

The above table 3 shows that the highest transaction data comes from year 2009 with a total of 846 transactions, whilst the lowest transaction data comes from year 2012 with a total of 159 transactions. For year 2012, the number of transaction data is lower because it only involves quarter 1 and quarter 2 of the year. Number of transaction property might not be completely updated as the transaction is still new.

For other years, such as year 2008 that comes with less transaction data, it might be due to the missing value in the data provided by VPSD. Any transaction data that comes with missing qualitative characteristics or missing house prices will be eliminated. This is in order to ensure the quality of data used in the modelling of house price index.

In this study, the construction of house price index is based on the transaction price of the property. The transaction price is used as the dependent variable which by it means that the property prices are depending on the characteristics of the property. The characteristics of the property used are physical and locational characteristics as have been mentioned earlier.

The ordinary least square (OLS) test assumptions

The model is tested according to the OLS assumptions as stated before. For heteroskedasticity test; Breusch-Pagan/Cook-Weisberg test, it shows that the model fail to reject the null hypothesis that is the constant variance. The result explained that the residuals are homoskedasticity. On the other hand, for the normality assumptions, the model's residuals are normally distributed. It is shown in figure 1 below:

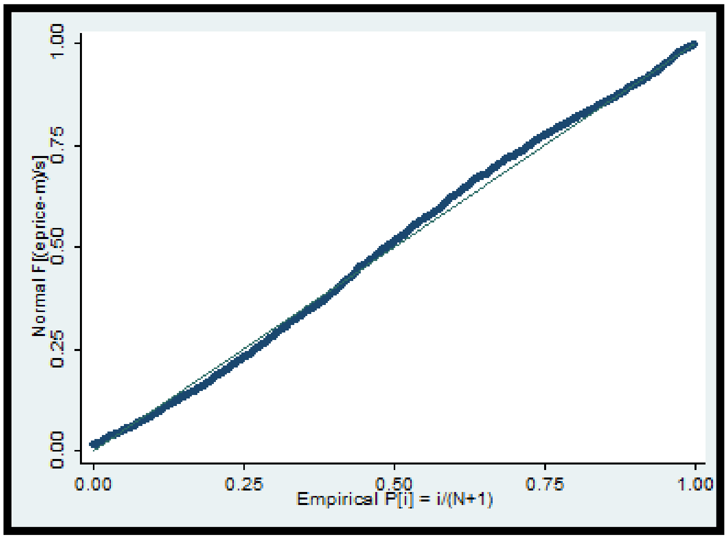

Figure 1: The normality assumption; residuals are normally distributed

Table 6 below shows the result of the hedonic regression that has been tested according to the OLS assumptions:

Table 6: The OLS estimates for double-storey terraced house market

Diagnostic test		
R^2	**0.6512**	
Adjusted R^2	**0.6479**	
F-statistic	**193.65**	
***p*-value (F-statistic)**	**0.0000**	
Breusch-Pagan/Cook-Weisberg	**0.1590**	
No. of observations	**3980**	
Variables	***Coefficient value***	***p-value***
Log (Lot area)	0.5820	0.000
Log (Built area)	0.5359	0.000
Log (Bedroom)	0.1305	0.000
Log (Distance)	-0.5038	0.000
Building age	-0.0153	0.000
Sub-district		
Batu (base)	-	-
Cheras	-0.2987	0.000
Kuala Lumpur	0.2066	0.000
Petaling	-0.1001	0.000
Setapak	-0.1185	0.000

*For the overall result of time dummy variable please refer appendix A.

Table 6 shows the result of the OLS diagnostic test. The value of R^2 is 0.6512. This indicates that about 65 percent of the log transaction price of house can be explained by its physical characteristic, locational characteristic and time-dummies variables. The *p*-value for all of the physical characteristic variables indicates that they are significant with the property prices. The coefficient value for those variables also show similar result produced by past studies. For lot area, built area and bedroom, the coefficient values are positive. This indicates that with the increasing value of those variables, the house prices will also increase. As an example, the bigger the built up area, the higher the house price.

On contrary, for distance and building age variable, the coefficient values are negative. For distance variable, the result follow the general hypothesis of the nearer the house to central business district (CBD) area, the higher the house price. Whilst for the building age variable, the value of house will decrease due to the depreciation that occurs to the house. From the hedonic regression result, it shows that the model is fit. It could be seen from the result produced by the diagnostic test. All of the independent variables used are uncorrelated with the residuals and for the Breusch-Pagan/Cook-Weisberg test; it indicates that the error term are homoskedasticity.

The hypothetical price index

The hypothetical price index for double-storey terraced house in Kuala Lumpur area is constructed by using the coefficient value produced by the hedonic regression. The coefficient value that will be used is the coefficient value of time-dummies variable. The construction of the hypothetical price index is carried out by taking the anti-log of the coefficient value for each quarter. The coefficient value for each quarter is available in appendix A. There are a total of 30 quarters which it begins from Quarter 1: 2005 to Quarter 2: 2012. Quarter 1: 2005 is the base period for the price index. It means that the price changes for the following quarter are based on that period.

Figure 2 below shows the hypothetical price index patterns produced by using the constructed model. The hypothetical price index is compared with the existing house price index, the MHPI produced by the VPSD.

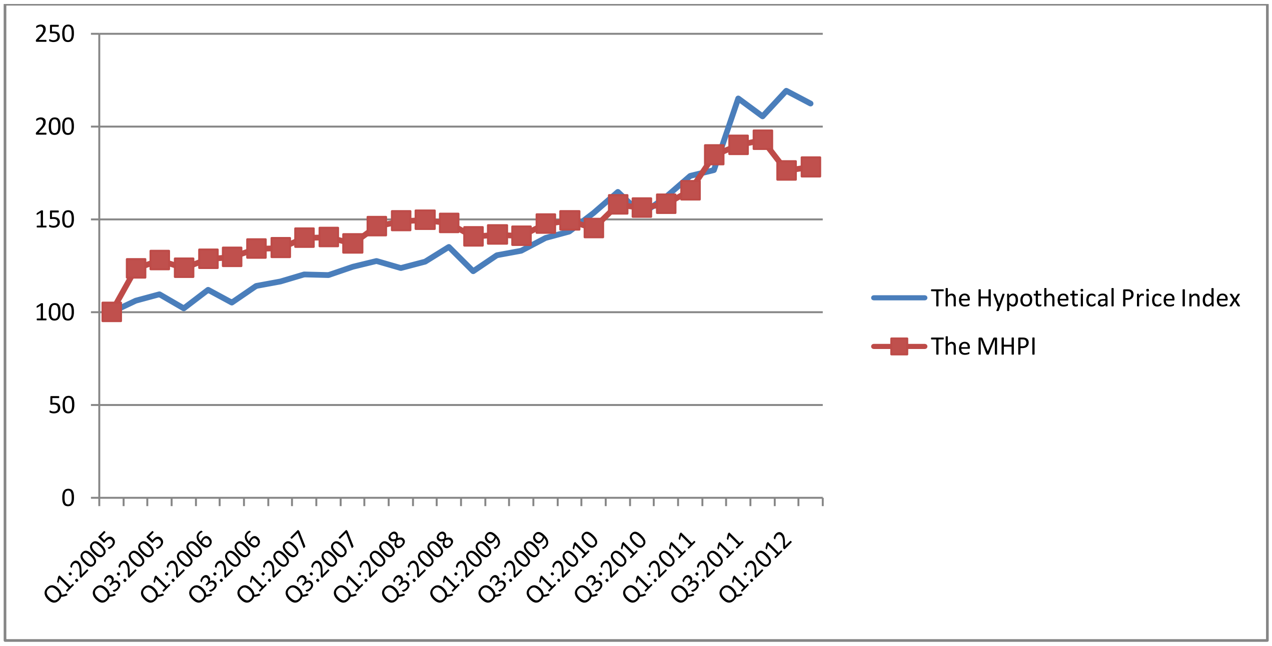

Figure 2: The constructed house price index versus the MHPI

From figure 2 above, it could be seen that the hypothetical price index is dissimilar with the MHPI produced by the VPSD at certain quarters. At Q2: 2006, Q2: 2007, Q3: 2007, Q3: 2008, Q2: 2009, Q4:2011, Q1: 2012 and Q2: 2012, there are some contradict values between both of the house price index. For instance, at Q6 which represent Q2: year 2006, the MHPI shows a declining pattern in the changes of house price. However, the hypothetical price index shows an increasing pattern in the house price changes. This situation might occur due to the dissimilarity of sample used. The sample used in this research mainly focus on the double-storey terraced house in order to construct the house price index. On contrary, sample used by VPSD include all type of terrace houses.

Besides, the dissimilarity between the two house price indices might arise due to the difference variables used in the construction of house price index. Some of the variables such as house type, building quality and tenure type that are used by the VPSD are not available to be used in this research. This is due to the limitation of data provided by NAPIC. Other than that, approach used by the VPSD to construct the MHPI is different with the one used in this research. This research adopts another approach of hedonic method; the time-variant index. The added of time-dummies variable in the hedonic equation helps to capture variation of house prices in each time period. On contrary, the MHPI is constructed by adopting the exact hedonic equation which it is compute for every single quarter. Thus, the difference between the hypothetical price index and the MHPI might occur due to the difference technique used.

Conclusion

House price index has become a requirement for various parties in order for them to monitor the house price changes from period to another period. In Malaysia, there is only one type of house price index that is available; the MHPI. Of this reason, this study has constructed a hypothetical price index so that parties that have interest in real estate market could have another alternative of house price index to be referred.

Method used to construct the hypothetical price index is same as method used to construct the MHPI. The only difference is the approach used which the MHPI is constructed using the exact hedonic equation whilst the hypothetical price index is constructed using time-variant index. The use of time-variant approach in the construction of the hypothetical price index is due to the limitation of transaction data used. The trends show from the hypothetical price index and the MHPI is quite similar. Only at certain quarters, the index is different. It might be due to the different sample of data used and also different variable used in constructing the model. However, the result still shows the similarity of trends between those two price indices; the hypothetical price index and the MHPI. Thus, it shows that the time-variant approach is also suitable to be used in the construction of house price index. It is hoped that the hypothetical price index that has been constructed in this study could be one of the alternative for parties that have interested in real estate market to monitor the house price changes.

References

Brunauer, W. A., Lang, S., Wechselberger, P., (2010). Additive hedonic regression models with spatial scaling factors: An application for rents in Vienna. *Journal of Real Estate Finance & Economics, 41(4), pp.390-411.*

Case, B. & Szymanoski, E. J. (1995). Precision in house price indices: Findings of a comparative study of house price index methods. *Journal of Housing Research, 6(3), pp. 483 - 96.*

Dorsey R. E., Hu H., Mayer W. J. & Wang H. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics, 19(2), pp.75-93.*

Kiel, K. A. & Zabel, J. A. (2008). Location, location, location: The 3L approach to house price determination. *Journal of Housing Economics, 17(2), pp. 175 – 190.*

Lehner, M. *Modelinh housing prices in Singapore applying spatial hedonic regression.* M.Sc. Thesis. 2011.

Lim, S. and M. Pavlou (2007), An improved national house price index using land registry data. *RICS Research Paper Series, 7(11).*

Longford, N. T. (Ed.) (2009). *A House Price Index Define in the Potential Outcomes Framework.* Barcelona, Spain: SNTL and UPF.

Norhaya K., Suriatini I., Hishamuddin, Ibrahim S., Rosadah S., (Ed.) (2008). *Modeling of the Property Market: The Malaysian Experience.* Universiti Teknologi Malaysia.

Osland L., (2010). An application of spatial econometrics in relation to hedonic house price modelling. *Journal of Real Estate Research, 32 (3), pp. 289-320* .

Ramanathan, R. (2002). *Introductory Econometrics with Applications.* Fifth edition. University of California-San Diego: South-Western Thomson Learning.

Rosen S., (2002). Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy, 82(1), pp. 34 - 55.*

Tan Y. K. (Ed.) (2011). *An Hedonic Model for House Prices in Malaysia.* Universiti Putra Malaysia.

Wallace N. E., (Ed.) (1996). *Hedonic-Based Price Indexes for Housing: Theory, Estimation and Index Construction.* University of California, Berkeley: FRBSF Economic Review.

Appendix A

Variables	*Coefficient value*	*p-value*
Time-dummies		
Q1 (base)	-	-
Q2	0.0612	0.058
Q3	0.0921	0.004
Q4	0.0219	0.524
Q5	0.1136	0.001
Q6	0.0517	0.113
Q7	0.1323	0.000
Q8	0.1519	0.000
Q9	0.1865	0.000
Q10	0.1821	0.000
Q11	0.2184	0.000
Q12	0.2432	0.000
Q13	0.2140	0.000
Q14	0.2417	0.000
Q15	0.3005	0.000
Q16	0.1999	0.000
Q17	0.2663	0.000
Q18	0.2847	0.000
Q19	0.3368	0.000
Q20	0.3607	0.000
Q21	0.4290	0.000
Q22	0.4999	0.000
Q23	0.4232	0.000
Q24	0.4819	0.000
Q25	0.5514	0.000
Q26	0.5685	0.000
Q27	0.7666	0.000
Q28	0.7200	0.000
Q29	0.7851	0.000
Q30	0.7526	0.000