THE CONSTRUCTION OF HOUSE PRICE INDEX:
MODELLING BY INCORPORATING ABSOLUTE LOCATION

NORSZAWANI AFIQAH BINTI ROSMERA

A thesis submitted in
fulfillment of the requirement for the award of the
Degree of Master of Science in Real Estate and Facilities Management

Faculty of Technology Management and Business
Universiti Tun Hussein Onn Malaysia

JUNE 2014
ABSTRACT

Price of residential property plays a very important role in the economy of developed countries. Price is an element for individual or household to decide on selling and buying properties, and to invest in the direct property market. In response to such a demand, most developed countries have produced a house price index to monitor price changes. In other countries, studies related to the construction of house price index are widely conducted. However, in Malaysian context, there is less attention regarding the construction of house price index. To fill in the gap, this research aims to construct a house price index by considering absolute location in modelling the house price index. Three models of house price index (HPI) are constructed. Method used in the construction of HPI is the hedonic method. The constructed HPI models are composed of different spatial elements as explanatory variables. Spatial elements included in this research are subdistricts dummy, distance of house to city centre, and coordinate of houses. The results show that the inclusion of property coordinate helps improve the model fitting and explains variation in house price. However, in terms of price index, the inclusion of coordinate of houses may not be important as it produces a similar movement pattern with price index that is composed of common locational variable. This research reveals that the common hedonic model that is composed of subdistricts dummy and distance to city centre as locational variable is sufficient to construct an HPI.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TITLE</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>xiv</td>
</tr>
</tbody>
</table>

## CHAPTER 1 INTRODUCTION

1.1 Introduction and research background  
1.2 Research issue  
1.3 Research aim  
1.4 Research questions  
1.5 Research objectives  
1.6 Research contribution  
1.7 Research scope  
1.8 Research design and methodology  
1.8.1 Research approach  
1.8.2 Data collection  
1.8.3 Data processing  
1.8.4 Technique of analysis  
1.8.5 Stage of study  
1.9 Thesis structure  

CHAPTER 2 THEORETICAL FOUNDATION OF REAL ESTATE

PRICE INDEX 12

2.1 Introduction 12
2.2 Overview of real estate price index 13
  2.2.1 The importance of price index in real estate market 14
2.3 Basis for house price index construction 16
  2.3.1 Appraisal-based index 16
  2.3.2 Transaction-based index 18
2.4 Method of real estate price index construction 20
2.5 Fundamental of hedonic price index 27
2.6 The taxonomy of hedonic method 32
  2.6.1 Time-dummy methods 32
  2.6.2 Hedonic imputation methods 33
2.7 Factors affecting real estate price index 34
2.8 Conclusion 38

CHAPTER 3 THE SPATIAL ELEMENT 39

3.1 Introduction 39
3.2 Definition of spatial 40
3.3 The needs to consider spatial elements 41
  3.3.1 Spatial effects in property data 41
    3.3.1.1 Spatial autocorrelation 42
    3.3.1.2 Spatial heterogeneity 45
3.4 The spatial element 46
  3.4.1 Spatial elements used as locational variables 46
3.5 Modelling house price using spatial element 50
3.6 Conclusion 59
CHAPTER 4 RESEARCH METHODOLOGY

4.1 Introduction 60
4.2 Research process 61
4.3 Characteristics of research method 63
4.4 Application of research method 64
4.4.1 Type of research data 65
4.5 The construction of real estate price index 66
4.5.1 Data background 66
4.5.2 Variable selection 67
4.5.2.1 Dependent variable 68
4.5.2.2 Independent variable 69
4.5.3 Method of analysis 74
4.5.3.1 OLS estimator 77
4.5.3.2 Functional form selection 79
4.5.4 Modelling the house price index 80
4.5.4.1 Application of time-variant approach 84
4.6 Conclusion 85

CHAPTER 5 THE CONSTRUCTED HOUSE PRICE INDEX ANALYSIS

5.1 Introduction 86
5.2 Analysis of models constructed 87
5.2.1 Descriptive analysis 90
5.2.2 Empirical analysis 92
5.3 Construction of HPI 97
5.4 Comparison of house price indices 100
5.4.1 Descriptive analysis comparison for the constructed HPI 100
5.4.2 The t-test analysis for the constructed house price indices 103
5.5 Conclusion 105
CHAPTER 6 RESEARCH CONCLUSION

6.1 Introduction 107
6.2 Results discussion 108
6.3 Justification of research findings 114
6.4 Research contribution 116
6.5 Limitation of the research 117
6.6 Recommendation for further research 119

REFERENCES 120

APPENDIX 127
# LIST OF TABLES

2.1 Methodology used in price index construction 25  
2.2 Variables used in determining the property prices 37  
3.1 Relationship between classes of spatial objects 40  
3.2 Summary of past studies on spatial elements 50  
4.1 Qualitative approach versus quantitative approach 64  
4.2 District and subdistricts 72  
4.3 Number of transaction data based on quarter 74  
5.1 Variables details for HPI Model I 88  
5.2 Variables details for HPI Model II 88  
5.3 Variables details for HPI Model III 89  
5.4 Descriptive statistics for double-storey terrace housing units used in the HPI constructions 90  
5.5 The empirical analysis result for HPI Model I 92  
5.6 The empirical analysis result for HPI Model II 94  
5.7 The empirical analysis result for HPI Model III 95  
5.8 The coefficient value and quarterly price index for HPI models 98  
5.9 Descriptive statistics for house price index level 101  
5.10 Descriptive statistics for house price index return 102  
5.11 Locational characteristics used as variables for each model 104  
5.12 Paired sample t-test between constructed HPI 104  
6.1 The advantages and disadvantages of house price index constructing methods 108
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Summary for the thesis elements</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Spatial autocorrelation between houses</td>
<td>43</td>
</tr>
<tr>
<td>3.2</td>
<td>Spatial heterogeneity in property data</td>
<td>45</td>
</tr>
<tr>
<td>4.1</td>
<td>Positive sign for estimated coefficient value</td>
<td>69</td>
</tr>
<tr>
<td>4.2</td>
<td>Negative sign for estimated coefficient value</td>
<td>69</td>
</tr>
<tr>
<td>4.3</td>
<td>Property characteristic used as explanatory variables</td>
<td>70</td>
</tr>
<tr>
<td>4.4</td>
<td>The exact and time-variant approach of hedonic method</td>
<td>76</td>
</tr>
<tr>
<td>4.5</td>
<td>Model used to construct the HPI</td>
<td>81</td>
</tr>
<tr>
<td>5.1</td>
<td>The constructed HPI</td>
<td>99</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

CBD - Central Business District
GIS - Geographic Information System
GWR - Geographically Weighted Regression
HPI - House Price Index
LAT - Latitude coordinate
LON - Longitude coordinate
MHPI - Malaysian House Price Index
MRA - Multiple Regression Analysis
NAPIC - National Property Information Centre
OLS - Ordinary Least Square
SPAR - Sales Price Appraisal Ratio
UTM - Universal Transverse Mercator
VPSD - Valuation and Property Services Department
LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Hedonic regression result for HPI Model I</td>
<td>127</td>
</tr>
<tr>
<td>B</td>
<td>Hedonic regression result for HPI Model II</td>
<td>129</td>
</tr>
<tr>
<td>C</td>
<td>Hedonic regression result for HPI Model III</td>
<td>131</td>
</tr>
<tr>
<td>D</td>
<td>Correlation analysis of model’s residuals and independent variables for HPI Model I</td>
<td>133</td>
</tr>
<tr>
<td>E</td>
<td>Result of t-test analysis for the constructed HPI models</td>
<td>134</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Introduction and research background

The advancement of technology provides opportunities to understand the dynamic of real estate market. Continuous effort can be advanced with the objective to produce an accurate real estate price indicator. Housing price index, which is an example of real estate price indicator, is an important tool to those involved in real estate market, including investors, financial institution, researchers, and developers. These people need such an index for a specific purpose (Gourieroux & Laferrere, 2009). In addition to that, policy maker also relies on housing price index prior to formulation of economics as the housing market contributes significantly to nation’s GDP.

According to Longford (2009), price index has been used to benchmark and monitor equity investment such as investment returns by property owners. Similarly, price index is also used by investors to make comparison with other alternatives. Netzell (2010) notes that the real estate price index is used to monitor the real estate cycle and the relationship between real estate markets. Thus, price indices particularly for residential market are very important to those who have interest in real estate market including financial services as well as the government.

The purpose of house price index is to provide an overview on the changes of house prices over time. For example, Longford (2009) posits that the house price index play a significant role in individual’s decision whether to buy or to sell a property. Other than that, house price index is also one of the important factors for policy makers who rely on property price signals to make a decision. As noted by
Lum (2004), real estate contributes to the nation’s wealth, and has become one of the investment media.

Due to the importance of housing market to the economy, many countries have produced their own residential price indices to monitor changes in housing prices. Such countries include the UK and the US. In these countries, the house price index has been established for the past 40 years. In Malaysia, an effort to produce the house price index started in 1993 by the National Property Information Centre (NAPIC), while the production of the house price index began in 1997 (Kamaruddin et al., 2008). Known as Malaysian House Price Index (MHPI), the house price index was introduced by the Valuation and Property Services Department (VPSD).

Tan (2011) explains that the construction of MHPI comprises price indices for 13 states and 2 federal territories. The index is established to monitor the changes of real estate price from one period to another, and to assist the formulation of economic policy (Kamaruddin et al., 2008). The approach used to construct the MHPI is the hedonic method. This method of price index construction has been used widely in the US and the UK (Kamaruddin et al., 2008; Bourassa, Hoesli & Sun, 2006).

As real estate price indices are important, they should possess some quality in order to provide an accurate measurement of house price movements (Bourassa et al., 2006). Recently, there are growing interests in improving methodological aspect of real estate price index. Studies such as by Quigley (1995) and Vries et al. (2009) have used various techniques to develop reliable real estate price indices.

A few methods can be used to construct the price indices, and one that is widely used is the median price index. However, this method exhibits bias in its result due to the heterogeneous nature of residential property. Therefore, methods that could control different qualities of housing characteristics are needed to ensure the reliability of real estate price indices. Examples of methods that can handle the heterogeneity of real estate market are hedonic method, repeat sales method, and hybrid method (Case & Szymanoski, 1995). Among these three methods, the hedonic method has attracted particular attention and wide research.

The hedonic method is based on the hedonic hypothesis that products or goods are valued according to their respective characteristics (Rosen, 1974). In real estate market, house is an example of a product, and its characteristics can be
categorised into three parts namely structure, accessibility, and neighbourhood (Malpezzi, 2002).

Longford (2009) notes that repeat sales method can only be applied if the property has been transacted more than once. Practically, this approach is difficult to employ due to the thinness of the property market in short and medium terms.

Finally, hybrid method is the method that combines two equations of a single transacted property and also property that is transacted twice and more (Quigley, 1995). To apply this method in the construction of real estate price index, the sample data of single and multiple sales are combined. By using the repeat sales sample data, the parameters are then determined.

1.2 Research issue

As noted earlier, the house price index is an important element in policy formulation and decision making. The price index especially the house price index helps monitor the price changes of real estate market. Most developed countries such as the UK and the US have more than one house price index. In the UK, house price indices such as Land Registry House Price Index, Halifax House Price Index, Nationwide House Price Index, Investment Property Databank (IPD) Property Index, and The Office for National Statistics (ONS) House Price Index have long existed. The most prominent house price index is the one produced by the IPD, which has published UK’s first reliable index of investment property performance. The price index produced by the institution is also available monthly, quarterly, and annually.

Nevertheless, such transparency in real estate market remains opaque in Malaysia. There is only one house price index that is publicly available, namely the MHPI. This index is constructed based on the hedonic method. Kamaruddin et al. (2008) state that the samples of residential transaction used in the construction of MHPI were categorised according to district, type of house, location and condition.

In addition, there is lack of studies on the construction of house price index in Malaysia. Studies that relate to the property market modelling in Malaysia include the hedonic pricing model for Penang condominium market (Chau & Chin, 2002), integration of hedonic pricing models with GIS (Ismail, 2006), and property mass valuation model (Jahanshiri, Buyong & Shariff, 2011).
It is crucial to construct a reliable price index that can provide an accurate indicator of real estate price trends. A number of difficulties in constructing real estate price index have been examined. The difficulties are mainly due to the characteristics of the real estate market itself, which is heterogeneous and has infrequent transaction. According to Netzell (2010), the heterogeneity of real estate market is due to the characteristics of the individual property itself such as size, age, and distance to a specific location. Next, the infrequent transaction is due to few observable properties in a given period. The few observable properties are usually referred to the commercial property. On the contrary, residential property normally has higher number of transactions and they are relatively comparable (Netzell, 2010).

Past studies have adopted various methods in constructing house price index. As noted by Nagaraja, Brown and Wachter (2010), one of the methods is by taking the summaries of property prices that comprise means or medians of particular period. Similar method has also been employed by the Australian Bureau of Statistics to construct house price index (Nagaraja et al., 2010). Other established methods for real estate price index construction are hedonic method and repeat sales method. In this research, hedonic method was employed due to its ability to overcome the limitation of the median price and repeat sales method (Dorsey et al., 2010).

The hedonic method models property prices by its respective characteristics. It was pioneered by Griliches in 1961. Since then, the method has been used to construct price indices for automobiles, computers, and other products. The early application of hedonic method in house price can be seen in studies done by Rosen (1974) and Goodman (1978). In modelling housing price, hedonic method regresses the house price on a vector of its characteristics.

According to Bourassa et al. (2006), there are two approaches to constructing real estate price index based on the hedonic method. First is by performing a separate regression for each period, and the estimated implicit prices should be applied to a standardised bundle of attributes. Second is by including time dummy variables as part of the model.

Moreover, variables to be included in hedonic method should be properly identified. The regression with respect to both functional form and independent variables should also be defined correctly to ensure that the model will meet all the requirements under the Ordinary Least Square (OLS) assumption, hence unbiased estimates of the house price indices could be obtained.
However, as noted by Long, Paez & Farber (2007), the application of hedonic method in modelling real estate price index results in spatial effects. These effects exist in the property data when one property influences others in terms of the market price, resulting in autocorrelation in a model outcome. In order to determine the accurate market price and to produce an unbiased estimator, spatial effects must be considered in modelling the price index.

Spatial autocorrelation, which can be seen in cross-sectional data (Anselin, 1999), is a result of spatial effect. As explained by Anselin (1999), spatial autocorrelation can be defined as the coincidence of value similarity with locational similarity. For example, houses located in same neighbourhood tend to have similar house prices. This price similarity could be positive or negative depending on the value and locational similarity. In real estate studies, positive spatial autocorrelation exists when properties that have similar values tend to cluster together. On the other hand, negative spatial autocorrelation is when the property values in a defined geographic area show dissimilar pattern.

Intuitively, properties that are located close to each other tend to have similar values because they exhibit similar characteristics. For instance, terrace houses with three bedrooms are usually clustered together in an area, while terrace houses with six bedrooms are usually clustered together in other area. This situation can be observed from the development concept in Malaysia where a particular type of houses is mainly clustered together in one housing scheme, and this further reinforces the spatial autocorrelation problems. As the distance between properties increases, spatial autocorrelation between properties is expected to decrease.

Therefore, to construct real estate price index by adopting hedonic method, spatial data must be considered. According to Anselin (1999), spatial econometrics is a subset of the econometrics method concerned with spatial aspects present in the cross-sectional data. Distance, location, and arrangement (topology) are the variables included in the spatial econometrics (Gerkman, 2010). In real estate market, location plays a very important role in contributing to the property prices. Even so, it is difficult to quantify the locational factor of property in quantitative form. The complexity to identify locational factors could bring problems to the validity of the Multiple Regression Analysis (MRA) model (Gallimore, Fletcher & Carter, 1996).

It is important to determine a precise geographical location for a property. The characteristics of a house that affect its price are spatially related in the form of
locational hierarchy (Kiel & Zabel, 2008). Recent studies show that location of a property is measured according to its distance with the nearest Central Business District (CBD), distance of the properties to the submarkets, and distance of the properties to the nearest public facilities (Gallimore et al., 1996).

However, Wilhelmsson (2002) notes that the inclusion of those locational characteristics still could not reject the hypothesis of no spatial effects. To encounter the spatial effect problem, this research used geocoded data that describe the precise location of a property. The availability of the exact location perhaps can contribute to determining the exact property prices. Studies done by Fik, Ling and Mulligan (2003), Gelfand et al. (2004), and Xu (2008) have incorporated geocoded data, i.e., coordinate \((x, y)\). These studies show that the variable is significant in house price modelling. Therefore, this research incorporated spatial elements in the construction of the house price index.

1.3 Research aim

The aim of this research is to model and subsequently construct the house price index by incorporating absolute location.

1.4 Research questions

The research questions are:

i. What are the methods applicable to be used in constructing house price index?

ii. How do spatial elements contribute in modelling a house price index?

iii. What is the model appropriate to construct a house price index?

1.5 Research objectives

i. To evaluate the applicability of existing methods for constructing a house price index.

ii. To investigate the significance of spatial element in modelling a house price index.
iii. To suggest an appropriate model for the construction of a house price index.

1.6 Research contribution

By achieving the research objectives, this research will contribute in the following ways:

i. Theoretical contribution
   a. This research helps identify essential elements of locational factor expectedly to improve the modelling of residential price index.
   b. This research extends the potential benefit of incorporating spatial elements in the modelling of residential price index.

ii. Practical contribution
   a. This research assists in improving decision making process by policy makers as a result of improving the accuracy level in price indices.
   b. This research helps related institutions especially banking institution in monitoring asset markets and in promoting financial stability.
   c. This research highlights to data provider, i.e., VPSD of the importance of spatial data in real estate market so that their data collection methods can be improved.

1.7 Research scope

This research is focused on the spatial effects that arise in the application of hedonic method to construct a real estate price index. An appropriate model of house price index is constructed by incorporating spatial elements in the model. GIS technique is employed in order to obtain the geocoded data of the observation points. Kuala Lumpur area is used as a case study for this research with the focus on the residential property.
1.8 Research design and methodology

1.8.1 Research approach

This research is based on quantitative techniques by employing secondary data. The secondary data include the transaction data of residential properties located in Kuala Lumpur from 2005 to 2012.

1.8.2 Data collection

Data used in this research are mainly from the secondary sources. The secondary data, which are the transaction data, are obtained from NAPIC. The data comprise transaction data from 2005 to 2012. A total of 3,200 transaction data covering Kuala Lumpur area are available to be used in the construction of house price index for this research.

1.8.3 Data processing

The transaction data obtained from NAPIC are analysed and processed using Stata II and SPSS software. Besides, Google earth software is used in order to obtain the geocoded data of houses involved in this research.

1.8.4 Techniques of analysis

To construct the real estate price index, hedonic method is employed in this research. This method, a special case of MRA, incorporates spatial elements in the price index model. It is based on the OLS technique that estimates the coefficient of time series attributes.

1.8.5 Stages of study

There are three main stages in this research. These stages are the theoretical, empirical, and evaluation. The first stage namely theoretical stage comprises literature review and theoretical overview. In this stage, literature and evidence
related to this research are examined and critically evaluated. The important areas reviewed are on the methods applied in the construction of real estate price index model. The review also focuses on the spatial element in order to solve for the spatial effects issue that exist in the hedonic price indices.

Second stage is the empirical stage. The use of spatial econometrics and Geographic Information Systems (GIS) is highlighted in order to address the issues of spatial autocorrelation in the present hedonic method. The empirical data are prepared prior to the identification of spatial element. The OLS model is used to estimate the unknown parameter in the multiple regression method.

Evaluation stage is the final stage. In this stage, the constructed house price index, which incorporates spatial elements, is evaluated. The constructed models are compared in order to identify the appropriate model for the construction of house price index.

1.9 Thesis structure

There are six chapters in this thesis. Each chapter discusses specific aspects of this research. The chapters are as follows:

Chapter 1 is the introductory chapter of the thesis. The chapter discusses preliminary, research background, research issue, research aims, research questions, research objectives, research contribution, research scope, research methodology, and thesis structure. This chapter is significant as it provides the framework for the whole study.

Chapter 2 discusses the literature review specifically on the construction of real estate price index. Basically, this chapter reviews the price index and the application of price index in real estate market. The method of constructing the real estate price index model is discussed in this chapter. It also examines the problems that arise in implementing hedonic price modelling to construct the real estate price index.

Chapter 3 is another part of literature review that discusses the spatial element of property. Spatial element used to describe the location of property and types of spatial effects present in the property data are discussed in this chapter. This discussion is very important in order to achieve the aim of this research, which is to incorporate spatial elements in the constructed house price index. This review helps
review and analyse the exact locational aspects that are suitable to describe the exact geographical location of property.

Chapter 4 discusses the method used to achieve the second and third objectives of this research. This chapter offers discussion on research process, characteristics of the research method, application of research method, and construction of real estate price index. To implement this research, secondary data are used. For the data processing, it involves the use of Stata II and SPSS statistical software to analyse and clean the collected data. The method used to construct the real estate price index modelling is the hedonic method.

Chapter 5 is the data analysis part. Prior to the data analysis, all of the collected data are cleaned and analysed using the statistical software. For the geocoded data, they are obtained from Google Earth software. All of the spatial data are stored in the software. The hedonic method, which is based on multiple regressions and OLS, is used to construct the real estate price indices. The constructed real estate price indices are compared between each other in order to find the model that is able to produce a better price index.

Finally, Chapter 6 summarises the whole study according to the three stages in this research. This chapter presents the findings of the study. The limitation of the study and potential areas for future research are also discussed in this chapter. Finally, this chapter is concluded by highlighting the contribution the research has made to knowledge. Figure 1.1 summarises the main elements of the thesis.
Theoretical Stage

Literature Review
a. Definition of price indices.
b. Methods to construct price indices.
c. Problems with the hedonic method in constructing the price index – the spatial effects problems.

Empirical Stage

Data Collection
a. Transaction data of the residential properties are obtained from NAPIC.
b. Journals.
c. Using Google earth software to obtain the geocoded data of the observation points.

Data Analysis
a. Data preparation, re-coding of variables and model analysis by using statistical software of Stata II and SPSS.
b. Use of hedonic price methods to construct the real estate price index modelling.

Evaluation Stage

Result and Discussion
a. Discussion on the findings and important result against the literature review in general.
b. The constructed house price indices are compared between each other to see which model produces a better index.
CHAPTER 2

THEORETICAL FOUNDATION OF REAL ESTATE PRICE INDEX

2.1 Introduction

This chapter discusses the theoretical aspect in developing the house price index. The discussion covers overview of real estate price index, basis for house price index construction, method of real estate price index construction, fundamentals of hedonic price index, taxonomy of hedonic method, and factors affecting real estate price index. Major concern of this research is to construct a real estate price index. Although there are many studies that have focused on this specific area, this research tries to fill in the gap by incorporating spatial attributes in the model. The research is conducted within the framework of factors that contribute to changes in house price and subsequently construct the price index.
2.2 Overview of real estate price index

House price index is an essential element in monitoring the changes in property prices. Residential property would normally serve as collateral for a mortgage. Slight changes in house price will affect banking equity level. According to Wallace (1996), housing market is important to an economy where there is high level of homeownership. Office for National Statistics (ONS) states that in the UK, level of homeownership increased from 23 percent in 1918 to 69 percent in 2001.

Besides, the US Census Bureau notes that homeownership level in the US has increased. In 1900, the level of homeownership was only 46.5 percent as compared to 2013 where the homeownership rate was 65 percent. Developed countries such as the US, the UK, Australia, and Hong Kong have their own house price indices with specific purpose to monitor changes in the real estate market. The level of homeownership basically depends on the easy-mortgage lending conditions. Whelan (2013) in his article notes that the US homeownership level soared to a peak of 69.2 percent in the middle of 2004 due to the easy-mortgage lending.

Other than that, in the UK, index for housing market has been existing since 1973 (Lim & Pavlou, 2007). In the beginning, the price index was produced by the mortgage providers, and recently it was produced by the government agency, the Land Registry. Nationwide House Price Index and Halifax House Price Index are house price indices produced by the mortgage provider. The house price indices are provided by Nationwide Building Society and Bank of Halifax. Their house price indices are based on hedonic method. On the other hand, price index produced by the Land Registry is known as the Land Registry House Price Index. The Land Registry House Price Index is the only index that is based on repeat sales method.

In the US, there are variations of house price indices produced by different public and private agencies. These indices include Freddie Mac House Price Index (FMHPI), S&P/Case-Shiller Indices, CoreLogic House Price Index, and IAS360 House Price Index. The FMHPI and S&P/Case-Shiller Indices are the main references of house price index. The FMHPI is introduced in the US with the purpose to measure house price inflation (Freddie Mac, 2012). The first series of the FMHPI began in January 1975. The index covers 50 states and 367 metropolitan statistical areas. On the other hand, the S&P/Case-Shiller Indices are designed to measure the
average changes in home prices in a particular geographic area. The price indices are calculated monthly and cover 20 major metropolitan areas.

In Malaysia, the only publicly available price index is the Malaysian House Price Index (MHPI). This index was introduced in 1997 by the Valuation and Property Services Department (VPSD) (Kamaruddin, 2008). The house price index was established due to the needs to monitor the changes in real estate price from a period to another period.

Generally, both the US and the UK adopt hedonic method to construct the house price indices. According to Ramalho (2011), the US Census Bureau has been using the hedonic method since 1968, whilst the application of hedonic method in constructing house price index in the UK has been seen since 1983. Similar method has been used by the VPSD in Malaysia to construct the MHPI (Kamaruddin, 2008).

2.2.1 The importance of price index in real estate market

This research focuses on house price index. A slight change in house price index will result in significant implication to the economy. The house price index is very important to the financial institution especially the banking sector that provides mortgage for housing. Changes in house price will significantly affect the banks’ balance sheet. This is the case with subprime loan crisis in the US that had subsequently affected financial system around the world. With house price index available, policy maker is able to detect a signal that shows distress in real estate market and act upon it.

The real estate market has become one of the factors that contribute wealth to certain countries. The contribution of the real estate market has made real estate price index as one of the significant elements to guide investors and policy makers in measuring its impact to the economy. Take the US as an example; 70 percent of the wealth portfolio for the average US household comes from the residential properties asset (Wallace, 1996). This is because residential real estate provides housing for families. It is the greatest source of wealth and savings for many of them. Besides, the commercial real estate, which includes apartment buildings, also creates spaces for jobs in retail, offices, and manufacturing. Therefore, real estate income provides a source of revenue for millions of people and it plays an important role in determining the wealth of a country.
As noted by Wallace (1996), residential market has its own cycle where it is prone to boom and bust. Thus, the cycle of housing market can be said as unpredictable. This cycle leads to the introduction of the volatility into the wealth of the average household, and subsequently to the market value of residential mortgage. The portfolio for this residential mortgage basically is held by the financial institution, pension fund, insurance companies, and individual investors.

According to Case and Wachter (2003), the first application of real estate price index is on monitoring the financial stability and asset markets. Market price is very essential especially for the financial institution. Market price can be defined as the highest price set by the owner to sell a property. With that price, there will be a willing buyer to buy the property. However, due to the changes that may arise in property market, such as changes in the availability of housing stock and changes in cost of financing, the property price might drastically change (Case & Wachter, 2003).

Other than that, the real estate price index is also increasingly important to the financial institution as the institution provides liquidity to the market through mortgage. Financial institution also provides loan for property buyers. For example, house is used as collateral when the buyers make a loan with bank. Herring and Wachter (1999) state that, when the price of real estate increases, it helps increase bank’s capital value as banks indirectly own this real estate. Hence, the increased capital value helps the banking institution to relax their lending standards. Therefore, banks tend to expose their institution to lending risk when property prices are high. The availability of a price index is essential as it could help banking institution forecast the property prices.

On the other hand, with respect to real estate bubble, it is an economic event that takes place in domestic and global real estate markets. The real estate bubble occurs periodically. According to Case and Wachter (2003), market bubble can be identified through monitoring changes in house price index. This is because the market bubble is marked by rapid increases in prices of real property such as housing until they reach unsustainable levels and then decline. Therefore, the property price index is very important for the financial institution in order for them to monitor the changes in real estate market.

The real estate price index is also important for the academic research. This is in order for academics to understand the real estate market deeper (Bourassa et al.,
2004). For example, real estate price index helps the academic research to analyse the property prices and the efficiency of the property market. The needs of real estate price index can be seen when a researcher carries out an analysis on housing affordability perspective. Bourassa et al. (2004) note that data such as the income factors of buyers, either these data affect the demand for property or not, are included in the property price index.

The application of real estate price index can also be seen in the perspective of property owners and investors. Property price index is used as a benchmark for these people to compare the average returns of property with other assets such as bonds and stocks (Netzell, 2010). Besides, real estate price index helps determine the correlation between asset classes. This is necessary when the risk for a portfolio is to be determined. When correlations between two assets are high, risk to invest in those assets will also be high. Hence, there will be no diversification benefits. This is because, when one asset turns down, it will affect other assets. In addition, property price index helps property owners and investors decide the appropriate time to invest in real estate market (Longford, 2009).

In summary, due to the significant contributions of the real estate market to the economy, an accurate real estate price index has become a need for various parties such as financial institution, construction institution, real estate institution, academics, property owners, and investors.

2.3 Basis for house price index construction

Generally, based on the type of housing data, there are two approaches in developing a house price index, namely appraisal-based index and the transaction-based index. Appraisal-based index is an index constructed using the valuation data derived from past appraisal exercise. On the other hand, the transaction-based index is constructed using the actual transaction data.

2.3.1 Appraisal-based index

The appraisal-based index is a type of index constructed using sample data of the assessed property. Netzell (2010) notes that the appraisal-based index is constructed by revaluing the similar sample of properties in each time period. The appraisal-
based index thus helps avoid heterogeneity issues that normally appear when different samples of properties are used to construct the price indices. Nevertheless, properties derived from similar sample may also have gone through some changes in quality and attributes over time. Therefore, these changes must be considered as they would affect property prices.

Using valuation data to construct a price index highly depends on the nature and quality of valuations (Netzell, 2010). There are several issues that have been discussed in adopting appraisal-based index. Some of the issues are regarding the smoothing effect and the accuracy of price determined during the valuation process (Fisher, Miles & Webb, 1999; Geltner, Macgregor & Schwann, 2003).

Smoothing effect is related to liquidity (Geltner et al., 2003). The liquidity can be defined as the volume of trading in an asset market. As liquidity increases, information on current transactions relevant to price determination also increases. When smoothing effect occurs, it will cause a series of price that demonstrates a stable trend across time. Basically, this stable trend is unable to precisely draw the actual market price movement and its volatility. Thus, it will draw to the wrong conclusion.

Basically, smoothing effect arises during valuation exercise. Geltner et al. (2003) note that when valuers make an appraisal, they tend to use a weighted average of the simultaneous information and historical appraisals. A valuer will make a comparison between the subject properties with a similar property’s characteristics. Value for the subject property will then be estimated from the comparable properties (Netzell, 2010). As an example, to determine the value for house A, a valuer will find comparable evidences with similar characteristics and located in the same housing scheme as house A. Then, the value for house A is derived from the value of the comparable evidences and some adjustments are made.

The appraised value index may differ from the market price and fundamental value (Geltner et al., 2003). Instead of showing the actual market price movements, the appraised value index will show the appraisal price movements. This basically is the effect of employing weighted average during appraisal time. If the appraisal value is used for the individual property level, it may be an optimum updating process. However, when the optimal appraisals are aggregated as an index, they should not be optimum for the index. Thus, the smoothing effect actually arises due to the comparison method applied by the valuers.
Basically, when applying the comparison method, the comparable evidences should be identical with the valued property (Longford, 2009). However, it is difficult to find a homogeneous property as properties are heterogeneous in nature. The heterogeneous characteristics of properties can be seen from a broad aspect. These characteristics include the physical characteristics, location, environment, and neighbourhood. Each of the characteristics contributes to different property prices. The heterogeneity nature of property leads to the issue of having a perfect comparable in determining the property price. Data obtained by valuers normally contain noise and these data might not be up-to-date (Netzell, 2010).

According to Netzell (2010), there are two solutions on how to derive an unbiased price index from appraisal-based index, namely zero-correlation method and reverse engineering method. The zero-autocorrelation method is based on the assumption that property markets’ return is unpredictable. On the other hand, reverse engineering method is based on estimated value of a parameter. This method represents weight for the current information related to the past information.

2.3.2 Transaction-based index

Transaction-based index is an index that relies on the actual transacted price of a property. The transaction prices basically are the result of negotiations between buyers and sellers. The construction of a transaction-based index requires large sample of data (Netzell, 2010). As a result, it is a challenging task to construct this specific index. Some property such as the commercial property only comes with few numbers of transactions. This is in contrast with residential property that consists of a large number of transaction data. Due to the requirement of numerous transaction data, only residential price index is suitable to be constructed. Nonetheless, it will lead to the heterogeneity problem in producing a reliable house price index.

Heterogeneity of property characteristics needs to be controlled in order to obtain a reliable index. Differences between transaction prices of one property to another exist due to differences in property characteristics as well as the transaction time. According to Netzell (2010), heterogeneity will lead to noise in the observed transaction prices and bias in an index. Two main factors that lead to bias are as follows: (1) property characteristics may change over time; and (2) different properties are transacted at different time.
The former factor basically focuses on the improvements made to the property characteristics or the whole property market. If there is any improvement made to the properties’ technical amenities across a whole market, changes in price should be observed due to quality improvement. Second, type of properties transacted should be observed properly. High-quality properties are normally transacted in certain time period, and this situation may lead one to believe that prices are raised more than they actually have during this period.

Literature addresses some solutions to overcome the heterogeneity issues. There are two methods to construct the transaction-based index, namely repeat sales method and hedonic method. Both methods have the advantages and disadvantages. The availability of the transaction data is very crucial to implement these two methods in the construction of transaction-based index. First method that helps avoid the heterogeneity problem in transaction-based index is repeat sales method. According to Netzell (2010), this method was first introduced by Bailey, Muth and Nourse in 1963. This method produces an index that compares the price of houses that have been transacted twice and more during the period of constructed index.

Although this method helps avoid the heterogeneity problem, there are a few problems with it. Netzell (2010) notes that this method requires a large number of transaction data, and therefore is not feasible for property markets. Only properties that have been transacted twice and more can be used for the index construction. Second, this method does not consider the changes that occur on the property such as any depreciation and renovations. Finally, the method only focuses on constructing index for properties that are transacted often, thus the properties might not be representative of the population.

Another method that helps control the heterogeneity problem is hedonic method. The basic of hedonic method in a property is viewed as a composite good. According to Netzell (2010), buying a property is like buying a set of goods. By using hedonic method, it aims to compute the marginal contribution of each good on the value of the composite goods. This value is obtained by regressing the transaction price of the property with its characteristics. The price level in different periods can be captured by including time-dummies in the regression, while the property characteristics help control for the heterogeneity.
2.4 Method of real estate price index construction

A few methods are available to construct house price index. These methods include hedonic method, repeat sales method, hybrid method, and SPAR method. Among these four methods, hedonic method has dominated the literature. It is widely used due to its ability to determine the house prices based on its characteristics.

*Repeat sales method*

Repeat sales method is one of the methods that is widely used to construct a house price index. According to Longford (2009), the method was first introduced by Bailey, Muth and Nourse in 1963. In 1987, improvements were made based on the suggestion from Case and Shiller. Fundamentally, this method is a generalisation of the old chain index. As noted by Lum (2004), the change in the price of an asset that has been sold twice is a quality-adjusted price change. A sample of the price changes can be used to construct a quality-adjusted price index. When a property has been sold twice, the difference of the log-price would be defined as the outcome variables.

All available methods to construct real estate price index have their benefits. The practical benefit that can be obtained by using repeat sales method is in terms of the requirement for the covariates. Longford (2009) notes that this method does not require any covariate, which can be defined as a secondary variable that affects the relationship between the dependent variable of price and the existing independent variables of physical characteristics of the property. Although some additional data can be used in the regression, this method does not require those data in its application. Lum (2004) supports this argument based on the fact that this method only relies on the sale price of houses and the date of sale to construct an index.

However, repeat sales method has its limitation to be used in the construction of house price index. This method requires a large number of transactions data. It is difficult to acquire such an amount of data for houses that have been sold twice and more. Dorsey et al. (2010) note that in Los Angeles County, for a three-year period between 2003 and 2006, it can be seen that 80 percent of the total transactions are not repeat sales. Moreover, Quigley (1995) supports that the repeat sales method might not be able to be used for housing market in short or medium run due to the thinness of data for houses that have been sold more than once. This lack of data will lead to sample bias in constructing house price index (Booth & Marcato, 2004).
Booth and Marcato (2004) suggest a sampling procedure to overcome shortage in transaction data. This method assumes that a property that has been transacted at different times can be used to construct the index. To implement this method, homogeneous properties need to be used as the samples in constructing the index. However, it is difficult to find homogeneous properties as there will be slight differences among the properties. Tan (2011) supports that no two houses are alike due to their wide variety of locational and physical attributes.

The repeat sales method does not have the ability to control the changes in property characteristics and their implicit prices over time. The quality of property might change overtime especially when the interval period between transactions is long (Lum, 2004). According to Longford (2010), a property transacted after ten years will normally have change significantly.

The changes in property conditions might arise due to some reasons such as improvements made to the property, its maintenance, and changes in the surrounding environment of the property. Those changes could be negative or positive. For instance, previously there is no development that can be seen around the housing scheme, but after certain periods, it is surrounded by a lot of amenities such as school and shopping complex. Therefore, price for houses located in the housing scheme will increase as compared to its previous price.

**Hybrid method**

The hybrid method was suggested by Case and Quigley in 1991. This method combines two equations of a transacted property and property that has been transacted twice and more. In other words, to construct the price index based on hybrid method, a single sale and a multiple sales of property are combined in its computation. The multiple sales used in this method provide important information to estimate the constructed model.

According to Quigley (1991), information obtained from the multiple sales is related to the characteristics of observed property. The characteristics of property can be differentiated from each other. Other than that, the characteristics of property also help improve the price index, \( P_t \) and the parameters, \( \beta \) efficiently. The information can be obtained by allowing the estimation of variance-covariance matrix.

For a single sale, it does not have the ability to differentiate the measured characteristics of property with the unmeasured one. It is different from the multiple
sales as it does not allow the idiosyncratic elements to be estimated. Besides, it does not measure the variance as a function in period between sales. Quigley (1991) notes that price index based on multiple sales data is more efficient than price index based on single sale data. However, the size of data for multiple sales is smaller as compared to size of data for single sale. This is one of the limitations in applying this method for constructing price index.

According to Quigley (1991), to apply the hybrid model in constructing the price index, the data of single and multiple sales are combined. Hence, by using the repeat sales sample data, the parameter can be determined. The log sale price is regressed with the log for housing characteristics, i.e., a set of dummy variables that represents time and a set of dummy variables for the involved properties.

The hybrid method has the ability to use all types of data regardless the number of the property being transacted. This method also benefits from the repeat sales data as it is able to compare the unmeasured attributes to estimate the variance and covariance errors. On the other hand, by having the single sales data, it contributes in estimating the parameters \( P_t \) and \( \beta \) more efficiently.

**Sale price appraisal ratio (SPAR) method**

SPAR is another alternative to construct house price index. According to Vries et al. (2009), the SPAR method has been applied in New Zealand since 1960 in order to measure the house price index. This method is similar to the repeat sales method. The data used are based on the matched pairs. General shortage of transaction data for the base period may arise as majority of houses sold on the observation period are not sold during the base period. Hence, the base period price of the houses is determined using appraisal of the houses.

Vries et al. (2009) note that the price indices can be value-weighted or equally weighted. Either to use the value-weighted or equally weighted is based on the index objective. For the value-weighted price index, it weights the index of an individual dwelling according to its base period prices. The value-weighted price index is constructed in order to see value changes of the housing stock. Based on the study done by Vries et al. (2009), the value-weighted arithmetic SPAR is as shown in the equation below:
\[ I_{\text{SPAR}, t} = \frac{\sum_{j=1}^{n_t} P_{j,t} W_{j,0} A_{j,0}}{\sum_{i=1}^{n_0} P_{i,0} W_{i,0}} = \frac{\sum_{j=1}^{n_t} W_{j,0} P_{j,t} A_{j,0}}{\sum_{i=1}^{n_0} W_{i,0} P_{i,0} A_{i,0}} \]

(2.1)

where \( P_{j,t} \) and \( P_{i,0} \) in the current period \( t \) and period 0 in which the houses are valued are the transaction price for houses \( j \) and \( i \). \( A_{j,0} \) and \( A_{i,0} \) are the respective appraisals. \( n_t \) and \( n_0 \) represent the number of houses sold at period \( t \) and 0. For the equation in the right hand side, it represents the basic idea for the value-weighted SPAR index. The numerator represents the price changes that are computed for each house sold in period \( t \). It is computed as the ratio of the actual transaction price and the appraisal. The value-weighted index could be explained from the expression of \( W_{j,0} = A_{j,0} / \sum_{j=1}^{n_t} A_{j,0} \).

The use of SPAR method could help control the changes in the quality mix of sample. In contrast, it does not control the quality changes for individual property. Therefore, Vries et al. (2009) state that there is a need to adjust the valuation in terms of improvements that have been made to the houses. Although this method could not help control the quality changes of a property, it helps control in the aspect of location based on the matched pair principle used as the base in this method.

**Hedonic method**

Hedonic method is a widely used method in the construction of real estate price index. It is applicable for a transaction-based data. The method was introduced by Griliches (1961), 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, prices of houses are valued according to their characteristics such as number of bedrooms, building area, and land area. This method can be used for different types of product. The value of each product is based on its own characteristics. Thus, this method helps to capture the heterogeneity of products.

Rosen (1974) also defines the hedonic price as an implicit price of its attributes that are determined by heterogeneity and 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 index. The regression analysis involves two types of variables namely dependent variable (i.e., price of a product) and independent variable (i.e., the characteristics of the product).

The hedonic method is mainly used in the construction of residential property price index. Studies done by Case et al. (1995), Wallace (1996), Gourieroux et al. (2009), Dorsey et al. (2010), and Tan (2011) use hedonic method in the construction of house price index. Most of the studies mentioned earlier compare hedonic method and other price index methods. Results produced from the studies show that hedonic method helps produce a better house price index than other methods.

There are some advantages in applying the hedonic method to construct house price index. First, the construction of house price index will be easier due to the availability of data. There is a large number of residential data because residential properties are often transacted. Unlike repeat sales method, hedonic method allows data of houses that have been transacted once to be used in the model. This is because repeat sales method relies on data of houses that have been transacted more than once. It is quite demanding to have those data because houses are normally transacted within five years. Dorsey et al. (2010) note that in Los Angeles County, approximately 80% of transaction data are not repeat sales for a three-year period.

Housing is normally treated as a heterogeneous good defined by a set of characteristics. The use of hedonic method helps embed the important attributes of a property in estimating its value (Tan, 2011). The important attributes of property comprise physical and locational characteristics of a property. For example, to determine the value of a house, one can consider the physical characteristics such as number of bedrooms, land area, building area, and locational characteristics such as distance to CBD and distance to the nearest facilities.

The characteristics of house are indexed by $j$ and the number of houses is produced by $n$. Based on the hedonic price function, the house price is defined. A mathematical relationship between the house prices and the house characteristics are embodied in them. This relationship can be seen from the regression equation in which each of the attributes is described independently. The regression analysis helps determine the attributes that provide high impact to the property values. Thus, one can differentiate which attribute actually contributes to the house prices.
REFERENCES


A Comparison of Index Construction Methods. *Journal of Real Estate 


from http://www.freddiemac.com/finance/fmhpi/


BB1v1_ch9.pdf.

Gelfand, A. E., Ecker M. D., Knight, J. R. &Sirmans C. F. (2004). The dynamics of 
location in home price. *Journal of Real Estate Finance and Economics*, 29(2), 
pp. 149-166.

Geltner, D., MacGregor, B. D. & Schwann G. M. (2003). Appraisal smoothing and 

Gerkman, L. (Ed.) (2010). *Topics in Spatial Econometrics With Applications to 
House Prices*, Helsingfors: Hanken School of Economics.

Gillen, K., Thibodeau, T., & Wachter, S. (2001). Anisotropic autocorrelation in 
5-30.

of Urban Economics* 5, pp. 471-484.


Government, General Series No. 73, National Bureau of Economic Research.

international perspective, *Occasional Papers* 58.


Theriault, M., Rosiers F. D. & Vandersmissen, M. H. (1999).GIS-based simulation of accessibility to enhance hedonic modeling and property value appraisal: An application to the Quebec City metropolitan area. Laval University, Quebec, Canada.


