METHODOLOGICAL ENHANCEMENT OF HEDONIC PRICE INDEX IN MALAYSIA

CHENG CHIN TIONG

UNIVERSITI TEKNOLOGI MALAYSIA
METHODOLOGICAL ENHANCEMENT OF HEDONIC PRICE INDEX IN MALAYSIA

CHENG CHIN TIONG

A thesis submitted in fulfilment of the requirement for the award of the degree of Doctor of Philosophy (Real Estate)

Faculty of Geoinformation and Real Estate
Universiti Teknologi Malaysia

JULY 2017
DEDICATION

“This thesis is dedicated to my parents for their endless support and encouragement”
ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my advisor Assoc Prof. Dr. Janice Lee Yim Mei and Prof. Dr. Hishamuddin Bin Mohd Ali for the continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Dr. Alias Bin Abdul Rahman, Prof. Dr. Md Nasir Bin Daud, and Assoc. Prof. Dr. Ibrahim @ Atan Bin Sipan, and Dr. Rohaya Binti Abdul Jalil for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

I gratefully acknowledge the funding received towards my PhD from the Mybrain 15 programme of Higher Education Ministry Malaysia. Moreover, I would like to thank Universiti Teknologi Malaysia and Faculty of Geoinformation and Real Estate who gave me access to the laboratory and research facilities. Without their precious support it would not be possible for me to conduct this research.

Last but not the least, I would like to thank my family and my friends for supporting me spiritually throughout writing this thesis in general.
ABSTRACT

Housing market is one of the largest asset sectors in Malaysia. Several parties are concerned about the performance of housing market including government, financial institution, market player and public. A proper computation of house price index using hedonic model is critical in monitoring housing performance to avoid biased estimation which due to complexity and misspecification of the model. This study has enhanced the conventional hedonic price index model through a combination of regression modelling and spatial model in order to yield more reliable house price index. The studied sample comprises 6,420 transactions of double storey terrace houses in Johor Bahru from year 2006 until year 2011. This study examined three types of regression modelling namely shrinkage, semiparametric and ordinary least squares method in constructing spatial hedonic price index model and conventional hedonic price index model respectively. An optimal hedonic price index model was ascertained according to the predictive power, accuracy and consistency test in the study. The result found that shrinkage estimator is robust when it comes to perform spatial hedonic price index model as compared to ordinary least squares and semiparametric method. Moreover, the house price index is further enhanced using temporal aggregation and seasonality analysis. The results show that seasonal adjusted monthly index is more effective in monitoring housing price performance. Therefore, shrinkage estimator, spatial hedonic model, temporal aggregation and seasonality analysis are important in enhancing the methodological aspects in constructing hedonic price index. In conclusion, the improved house price index can be used in formulating more effective housing policies and investment strategies.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td></td>
<td>xv</td>
</tr>
<tr>
<td>LIST OF APPENDIX</td>
<td></td>
<td>xvii</td>
</tr>
</tbody>
</table>

## 1 INTRODUCTION

1.1 Background of Study 1
1.2 Problem Statements 4
1.3 Research Questions 10

1.3.1 Research Objectives 10
1.4 Scope and Limitation of Study 11
1.5 Significance of Study 11
1.6 Research Framework 12
1.7 Thesis Structure 14
2  HOUSE PRICE INDEX CONSTRUCTION METHODS  16

2.1 Introduction  16
2.2 Price Index as Economy Indicator  16
2.3 House Price Index  17
2.4 Price Index Formula  19
  2.4.1 Laspeyres Price Index  19
  2.4.2 Paasche Price Index  20
  2.4.3 Fisher Ideal Index  20
2.5 Housing Index Construction Method  21
  2.5.1 Simple mean or Median Method  21
    2.5.1.1 Mean  21
    2.5.1.2 Median  22
  2.5.2 Stratification (Mix Adjustment)  22
  2.5.3 Repeat Sales Method  23
  2.5.4 Appraisal Based Method  25
  2.5.5 Hedonic Regression Method  25
    2.5.1.1 Hedonic Model  26
    2.5.1.2 Hedonic Time Variable Method  27
    2.5.1.3 Characteristic Price Approach  27
    2.5.1.4 Hedonic Imputation Method  28
  2.5.6 Strength and Weakness of House Price Index Construction Method  29
  2.5.7 Types of Hedonic Regression Methods  37
  2.5.8 Practical Issues in Hedonic Regression Methods  40
2.6 Index Construction Method Used By Official Index Provider  42
  2.6.1 Malaysian House Price Index (MHPI)  43
2.7 Hedonic Modeling  44
  2.7.1 Hedonic Variables of Malaysian Hedonic Price Index Model  51
2.8 Summary  52
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3</td>
<td>3.2</td>
<td>Hedonic House Price Index Improvement</td>
</tr>
<tr>
<td>3</td>
<td>3.3</td>
<td>Past Studies of Housing Index Improvement</td>
</tr>
<tr>
<td>3</td>
<td>3.4</td>
<td>Application of Spatial Econometric</td>
</tr>
<tr>
<td></td>
<td>3.4.1</td>
<td>Spatial Dependence</td>
</tr>
<tr>
<td></td>
<td>3.4.2</td>
<td>Spatial Heterogeneity</td>
</tr>
<tr>
<td></td>
<td>3.4.3</td>
<td>Spatial Autoregressive Model</td>
</tr>
<tr>
<td></td>
<td>3.4.3.1</td>
<td>First Order Spatial AR model</td>
</tr>
<tr>
<td></td>
<td>3.4.3.2</td>
<td>The Mixed Autoregressive Regressive Model</td>
</tr>
<tr>
<td></td>
<td>3.4.3.3</td>
<td>The Spatial Autoregressive Error Model</td>
</tr>
<tr>
<td></td>
<td>3.4.3.4</td>
<td>The Spatial Durbin Model</td>
</tr>
<tr>
<td></td>
<td>3.4.3.5</td>
<td>The General Spatial Model</td>
</tr>
<tr>
<td></td>
<td>3.4.4</td>
<td>Trend Surface Analysis and Spatial Correlation</td>
</tr>
<tr>
<td></td>
<td>3.4.1.1</td>
<td>The Construction of Spatial Hedonic Price Index Model Using Trend Surface Analysis</td>
</tr>
<tr>
<td>3</td>
<td>3.5</td>
<td>Multivariate Regression Analysis for Hedonic Model</td>
</tr>
<tr>
<td></td>
<td>3.5.1</td>
<td>The Unbiasness of OLS</td>
</tr>
<tr>
<td></td>
<td>3.5.2</td>
<td>The Variance of OLS Estimator</td>
</tr>
<tr>
<td>3</td>
<td>3.6</td>
<td>Semiparametric Regression</td>
</tr>
<tr>
<td></td>
<td>3.6.1</td>
<td>Application of Semiparametric Regression In Formulating Spatial Hedonic Model</td>
</tr>
<tr>
<td>3</td>
<td>3.7</td>
<td>Spatial Autoregressive Model, Trend Surface Analysis Model and Semiparametric Bivariate Smoothing Model for Constructing Spatial Hedonic Model</td>
</tr>
<tr>
<td>3</td>
<td>3.8</td>
<td>The Understanding of Bias and Variance</td>
</tr>
<tr>
<td>3</td>
<td>3.9</td>
<td>Shrinkage Estimation</td>
</tr>
<tr>
<td></td>
<td>3.9.1</td>
<td>Stein Rule Estimation</td>
</tr>
</tbody>
</table>
5.4 Predictive Power of Hedonic Price Index Model 126
5.5 Level of Index Revision 130
5.6 Accuracy of House Price Index 133
5.7 The Selection of Robust Hedonic Price Index Model 137
5.8 Temporal Aggregation Analysis 139
  5.8.1 House Price Index with Different Level Level of Temporal Aggregation 139
  5.8.2 Analysis of Variance (ANOVA) 141
  5.8.3 Performance of House Price Index 142
5.9 Seasonality of House Price Index 144
  5.9.1 Seasonality Analysis 145
5.10 Seasonal Adjustment 146
5.11 Conclusion 149

6 CONCLUSION AND RECOMMENDATION 151
  6.1 Introduction 151
  6.2 Research Summary and Findings 151
  6.3 Research Contribution 153
  6.4 Limitation and Further Research 158
  6.5 Conclusion 160

REFERENCES 161

Appendices A-N 173-308
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary on House Price Index Construction Method</td>
<td>34</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary Hedonic Regression Methods</td>
<td>37</td>
</tr>
<tr>
<td>2.3</td>
<td>Index Construction Method Used By Several Countries</td>
<td>43</td>
</tr>
<tr>
<td>2.4</td>
<td>Hedonic House Price Index Model</td>
<td>45</td>
</tr>
<tr>
<td>2.5</td>
<td>Past Development of Hedonic Models in Malaysia</td>
<td>49</td>
</tr>
<tr>
<td>4.1</td>
<td>Types of Data</td>
<td>100</td>
</tr>
<tr>
<td>4.2</td>
<td>Selected Variables</td>
<td>101</td>
</tr>
<tr>
<td>4.3</td>
<td>Summary of Final Dataset</td>
<td>103</td>
</tr>
<tr>
<td>4.4</td>
<td>Description on Data Quantification</td>
<td>106</td>
</tr>
<tr>
<td>5.1</td>
<td>Regression Analysis of Spatial Hedonic Model</td>
<td>120</td>
</tr>
<tr>
<td>5.2</td>
<td>Descriptive Statistics of Hedonic Price Index</td>
<td>123</td>
</tr>
<tr>
<td>5.3</td>
<td>The MAPE of Hedonic Price Index Model</td>
<td>127</td>
</tr>
<tr>
<td>5.4</td>
<td>The Level of Index Revision of Hedonic Price Index Model</td>
<td>131</td>
</tr>
<tr>
<td>5.5</td>
<td>The Accuracy of Hedonic Price Index Model</td>
<td>131</td>
</tr>
<tr>
<td>5.6</td>
<td>Ranking for Each Hedonic Price Index Model</td>
<td>138</td>
</tr>
<tr>
<td>5.7</td>
<td>ANOVA of Temporal Aggregated House Price Indices</td>
<td>142</td>
</tr>
<tr>
<td>5.8</td>
<td>Housing Performance of Temporal Aggregated House Price Indices</td>
<td>143</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.9</td>
<td>Regression Analysis of Seasonality Model</td>
<td>145</td>
</tr>
<tr>
<td>5.10</td>
<td>Comparison Between AHPI and NAHPI</td>
<td>148</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Study Flow Chart</td>
<td>14</td>
</tr>
<tr>
<td>3.1</td>
<td>Trend Surface of Different Order</td>
<td>67</td>
</tr>
<tr>
<td>3.2</td>
<td>Shrinkage Intensity and Mean Squared Error</td>
<td>80</td>
</tr>
<tr>
<td>3.3</td>
<td>Optimization Methodology Flow Chart</td>
<td>94</td>
</tr>
<tr>
<td>4.1</td>
<td>Case Study</td>
<td>92</td>
</tr>
<tr>
<td>5.1</td>
<td>House Price Index of Double Storey Terrace House</td>
<td>124</td>
</tr>
<tr>
<td>5.2</td>
<td>Four Types of Temporal Aggregated House Price Index</td>
<td>136</td>
</tr>
<tr>
<td>5.3</td>
<td>Comparison Between Adjusted and Non Adjusted House Price Index</td>
<td>147</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

NAPIC - National Property Information Centre
MHPI - Malaysian House Price Index
OLS - Ordinary Least Square
SR - Stein Rule
ANOVA - Analysis of Variance
MAPE - Mean Absolute Percentage Error
VPSD - Valuation and Property Services Department
PCA - Principal Component Analysis
WRS - Weighted Repeat Sale
DTH - Hedonic Time Dummy
HIM - Hedonic Imputation
RLS - Restricted Least Squares
STAR - Spatiotemporal Autoregressive Model
SHM - Spatial Hedonic Model Performed by Ordinary Least Square
SRSHM - Spatial Hedonic Model Performed by Shrinkage Estimation
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPM</td>
<td>Spatial Hedonic Model Performed by Semiparametric Regression</td>
</tr>
<tr>
<td>OLS</td>
<td>Simple Hedonic Model Performed by Ordinary Least Square</td>
</tr>
<tr>
<td>SROLS</td>
<td>Simple Hedonic Model Performed by Shrinkage Estimation</td>
</tr>
<tr>
<td>SPOLS</td>
<td>Simple Hedonic Model Performed by Semiparametric Estimation</td>
</tr>
</tbody>
</table>
**LIST OF APPENDICES**

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Semiparametric Regression Analysis</td>
<td>176</td>
</tr>
<tr>
<td>B</td>
<td>Eigen Value of Empirical vs Full Shrinkage Model</td>
<td>181</td>
</tr>
<tr>
<td>C</td>
<td>Analysis of MAPE</td>
<td>185</td>
</tr>
<tr>
<td>D</td>
<td>Analysis of Level of Index Revision</td>
<td>189</td>
</tr>
<tr>
<td>E</td>
<td>Analysis Breakdown of Index Revision</td>
<td>206</td>
</tr>
<tr>
<td>F</td>
<td>Analysis of Accuracy of House Price Index Model</td>
<td>209</td>
</tr>
<tr>
<td>G</td>
<td>Index Model Evaluation</td>
<td>213</td>
</tr>
<tr>
<td>H</td>
<td>Shrinkage Intensity of Shrinkage Models</td>
<td>249</td>
</tr>
<tr>
<td>I</td>
<td>Knots,Df And Spar Of Semiparametric Models</td>
<td>255</td>
</tr>
<tr>
<td>J</td>
<td>T-Test For MAPE</td>
<td>277</td>
</tr>
<tr>
<td>K</td>
<td>T-Test for Level of Index Revision</td>
<td>281</td>
</tr>
<tr>
<td>L</td>
<td>T-Test for Accuracy of House Price Index</td>
<td>284</td>
</tr>
<tr>
<td>M</td>
<td>Seasonality Analysis</td>
<td>288</td>
</tr>
<tr>
<td>N</td>
<td>Multicollinearity Test for SHM</td>
<td>292</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Background of Study

Housing market is one of the largest asset sectors in Malaysia. Several parties are concerned about the performance of housing market including government, financial institution, market player and public. This is because the downtrend or uptrend of housing performance will directly affect the decision making by those parties. Therefore, house price index is an important indicator to measure housing performance.

Rapid growth in property sector leads to the transformation of physical property investment into financial products such as housing derivative or home equity insurance. It stimulates the demand for more reliable, accurate, stable and warrant house price index to measure performance (Venter, 2008; Ong and Ng, 2009). Clapham et al. (2006) also mentioned that instability in measuring price movement is unfavourable to the development of markets for financial assets based on housing index. Therefore, a reliable house price index is needed to assist such development.

The heterogeneous characteristic of transacted properties in terms of locality and structural characteristic is a challenge in developing a house price index. Therefore, a quality adjustment should be considered in price index development. Otherwise, the housing index could be misleading. According to de Haan and Diewert (2011), the popular methods to develop house price index are repeat sales
method, hedonic regression method, mean/median mix adjustment method and appraisal base method. Among these methods, repeat sales method and hedonic regression method are most popular in constructing a quality adjusted price index. The official house price index which adopted repeat sales method is Case-Shiller Residential Price Index from United States. Meanwhile, Malaysian House Price Index (Malaysia), Centa-City Index (Hong Kong), and Halifax House Price Index (U.K) are using hedonic regression method to construct quality adjusted price index.

In comparing both repeat sales and hedonic regression the latter will be more appropriate to be applied in Malaysia. The repeat sales method requires same property to be transacted at least twice. In order to construct the price index it needs more property transaction. Moreover, according to Clapham et al. (2006) hedonic model is suitable to assist the development of housing derivatives index compared to repeated sales and median index because it is less prone to index revision. Since Malaysia is a developing country, it does not produce a big volume of property transaction for each period of time. Therefore, the application of repeat sale method is rather difficult for Malaysia. Hedonic regression method is the most appropriate method for Malaysia to construct house price index. Hedonic regression method is more flexible than other methods in term of constructing price index for specific market because all historical property transaction information can be used to develop house price index.

Although hedonic model can be used to construct house price index easily, the econometric problems will affect the reliability of hedonic model. The econometric problems consist of omitted variables bias, multicolinearity, overfitting, and selection of functional form in a hedonic model. Therefore, the study towards optimization of hedonic model is crucial for the development of a reliable house price index for Malaysia. The application of spatial model, shrinkage and semiparametric estimators for improving hedonic price model had been studied in past research. However, it is lack of study in combining these applications to improve hedonic house price index model. As Hoffmann and Lorenz (2006) say, “the future will belong to hedonic indices”.
Apart from that, house price index can be represented in different temporal aggregation level such as monthly, quarterly, semi-annually and annually. However, higher level of temporal aggregation will lead to missing information on the volatility of house price. In past, house price index is assumed to be smooth and less risky because real estate is known as hedge against inflation (Hoesli, 1994). However, the market sentiment at disaggregated periods such as monthly could be different. It can yield the volatility of house price in short term. A low level of temporal aggregated price index will unveil more information for improving decision making of property market players. Malaysia house price index (MPHI) is constructed in quarterly basis therefore it could lead to temporal aggregation problem and short term volatility is neglected. As claimed by Ramirez (2012) disaggregated data and models can enhance the accuracy of time series forecasting. Thus, temporal disaggregated price index would be effective to assist the formulation of national economic policy.

Besides, a house price index moves in a cyclical trend cannot be considered as effect of macroeconomic or market forces. It is because the price change is due to some events, festivals or seasons. Every country consists of different culture and seasons that will impact on such trend. Therefore, seasonality analysis is necessary to be conducted before releasing the price index to public. It is important for policy maker to determine the actual movement of house price without the seasonal effect. The seasonal pattern of the Malaysian house price index will mislead the interpretation of house price movement. As mentioned by Harding et al. (2003) the seasonal pattern is due to the bargaining power of buyer and seller. Seasonality analysis is needed to unveil Malaysian property market behavior. The understanding of property market behavior will help seller and buyer in obtaining good deal.

Thus, an optimal methodology is important because it can yield a reliable house price index for decision making of various parties in real estate such as valuer, developer, policy maker, investor and public.
1.2 Problem Statements

Malaysian House Price Index (MHPI) is an official price index for Malaysia housing market since year 1997. The index construction method adopted by MHPI is hedonic characteristics method and it is using Laspeyres index formula for price index computation. Basically, it is a simple hedonic model that consists of housing attributes and neighborhood information. The hedonic price index model is estimated through ordinary least square (OLS). In addition, it is a quarterly house price index without seasonal adjustment.

The hedonic characteristics method is an unconstrained hedonic model which required two separate hedonic models to compute price index. According to Schwann (1998) this method needs at least 32 observations to compute a viable model. Therefore, the hedonic characteristics method will need at least 64 observations to compute a house price index. It is a major weakness of this method because it could hamper the development of house price index for thin market due to insufficiency of observations. It would make the prediction becomes inaccurate due to overfitting and multicollinear problems. Therefore, hedonic time dummy method (single hedonic price index model) is more efficient than hedonic characteristic method (two separate hedonic price index model) in computing hedonic price index model.

Moreover, Hill (2012) claimed that the incorporation of geospatial data in hedonic characteristics model is rather vague. It is because this model will require average attributes to be substituted into the model such as average land size and etc. However, the definition of average location is ambiguous since the housing location is the composition of geo-coordinates x and y. The average value of x and y is not equal to average location. The situation becomes worse if the location of base period model (0) and respective period model (t) are totally different. The computed price index could be misleading. Hence, hedonic time dummy method is more flexible than hedonic characteristics in forming a spatial hedonic model using geospatial data. A spatial hedonic model is important to overcome the issue of spatial autocorrelation and spatial dependences as mentioned in previous research.
Therefore, hedonic time dummy is more suitable than hedonic characteristic method for constructing Malaysian house price index due to simplicity and flexibility. However, hedonic model will encounter econometric problems and it should be mitigated in order to yield a better model. In case of hedonic model, omitted variables bias is likely to happen if significant parameters are not included in a model. It can lead to large variance in the model and the predictive power becomes weak. According to previous research, the significant parameters for house price model are structural characteristics such as land area, built up area and number of bedroom. Besides, location is one of the important elements which affects the housing value due to maturity of neighbourhood. “Location!, Location!, Location!” which is often mentioned in real estate appraisal. A house in an urban area will definitely fetch higher value than a house in a suburban area. The location variables or geospatial data can be collected in the form of coordinates. If the geospatial data are not included in the model it would lead to spatial dependences and spatial heterogeneity problem and the predictive power of the hedonic model becomes poor.

Hill (2013) claimed that most of the researchers are using geospatial data to model housing market rather than developing housing index. The official house price indices that included geospatial data in hedonic price index model are limited such as FNC Residential Price Index and RPData-Rismark Daily Home Value Index. Furthermore, the adoption of geospatial data in hedonic price index model is given less attention by academic research (Schwann, 1998; Clapham et al., 2006; Wilhelmsson, 2009; Ericson et al., 2013). Although previous research (Can and Megbolugbe, 1997; Clapp, 2004; Sun et al., 2005; Liu, 2013) adopt various spatial hedonic models to compute house price index, the strength of spatial hedonic price index is not clearly highlight in their study. Moreover, their research are similar to the previous research of hedonic pricing model (Pace et al., 1998; Gelfand et al., 2003; Daria, 2007; Liu, 2013). Although spatial hedonic price index model outperforms simple hedonic price index model in term of predictive power, the differences between both indices are still questionable. As mentioned by Song and Wilhelmsson (2010), the relationship between predictive power of price index model and price index movement is unclear.
In Malaysia, it is lack of academic research in constructing house price index using spatial hedonic price index model (Afiqah et al., 2012; Lizam et al., 2013) as compared to hedonic pricing model (Ismail et al., 2009; Maimun, 2011; Ismail and Yusof, 2012). The main advantage of spatial hedonic price index model is still unclear. Malaysian house price index has tried to avoid spatial heterogeneity by using a defined neighborhood. However, Fik et al. (2003) claimed that it is hard to define the boundary with little knowledge of neighborhood. Thus, spatial model (using geocoordinates x,y as location variable) is better than aspatial model (using neighborhood as location variable).

The existence of collinearity problem in hedonic model can adversely affect the reliability of hedonic model and lead to overfitting issue. This problem will produce high value of R-squared. However, the predictive power is weak. Although hedonic model is flexible to include all quality adjusted elements, it will increase the complexity of the model. High complexity of model will induce extremely large error. It becomes worst if number of variables (p) is larger than sample size (n). Therefore, it will hamper the development of spatial hedonic model if the number of transacted property is thin. It is because inclusion of geospatial data will make the model becomes lengthy and complex especially a polynomial form of spatial hedonic model (Fik et al., 2003; Clapp, 2004). Generally, total error of model is composition of bias and variance. Although high complexity has less bias, it will increase the variance of the model simultaneously. Therefore, an optimum composition between bias and variance should be taken into consideration in computing a model. In order to calibrate an optimum spatial hedonic price index model, shrinkage estimation is recommended to be adopted in this study (Bao and Wan, 2007).

In previous study, Bao and Wan (2007) used stein rule estimation (SR) to perform a hedonic price model. SR estimator is a form of shrinkage estimation that employs a weighted average of OLS and restricted least square (RLS) estimator. The result found that SR estimation can overcome the collinear problem because it has enhanced the predictive mean square error (MSE) of house price model. In addition, a similar result is found in other related research (Knight and Hill, 1992; Knight and Hill, 1993; Namba and Ohtani, 2012). In past studies, most of the hedonic models
used to construct index through OLS estimation (Costello and Watkins, 2002; Clapham et al., 2006; Goh et al., 2012). Furthermore, Schafer and Strimmer (2005) and Opgen-Rhein and Strimmer (2007) found that shrinkage estimation (Ledoit_Wolf theorem) is applicable for bioinformatics model using small sample size (n) and larger number of variables (p). The result shows that shrinkage estimation outperforms OLS in term of mean square error. Therefore, it is useful for the development of hedonic price index model using small sample (n) and larger number of variables (p). The application of shrinkage estimation for strengthening the reliability of spatial hedonic price index model is deem to be feasible.

Moreover, functional form misspecification would affect on the accuracy of hedonic model in predicting housing index. According to Bao and Wan (2004) if wrong functional form is selected to compute hedonic model the validity of interpreting the estimated parameters will be highly questionable. Generally, semi-log is a famous functional form proposed by Box and Cox (1964) and widely used in constructing housing index (Goodman, 1978; Halvorsen and Pollakowski, 1981; Maurer et al., 2004; Clapham et al., 2006; Widlak and Tomczyk, 2010). Unfortunately, Coulson et al (2008) claimed that semi-log model could expose to the risk of functional form misspecification.

Breiman (2001) point out that semiparametric and non parametric approaches are representative of algorithmic modelling culture. Both methods are suitable for many hedonic modelling situations where incomplete knowledge prevents the exact a priori specification of non linear or non stationary components of functional form (Hannonen, 2006). Song and Wilhelmsson (2010) mentioned that more research are needed in order to gain a better insight into the art of specifying the functional form of hedonic equation.

Some past studies claimed that nonparametric or semiparametric method better than parametric models in term of out of sample prediction (Messe and Wallace, 1991; Pace, 1993; Pavlov, 2000; Clapp, 2004). However, as mentioned by Hill (2013) there is lack of discussion on nonparametric method in constructing housing index. Clapp (2004) has applied semiparametric in developing house price
index model. The result shown that, semiparametric methods reduce out of sample mean squared error by 11% as compared to the OLS. In Malaysia, due to lack of study in this area the application of semiparametric or non parametric in hedonic price index model is still questionable. Therefore, a semiparametric type of hedonic price index model is another possible study suggested by Bao and Wan (2004).

The predictive power is a common method to evaluate the reliability of house price model in most of the research. In fact, the relationship between predictive power of hedonic price index model and house price movement is rather vague. For example, Song and Wilhelmsson (2010) claimed that although the predictive power enhanced by spatial hedonic price index model, it does not affect the house price index of condominium significantly. Furthermore, some research emphasize the consistency of price index (Clapham et al., 2006; Baroni et al., 2008; Deng and Quigley, 2008) because the index revision can make the settlement of future contracts based on housing index becomes less precise. Hence, it is necessary to revisit the methods for evaluating hedonic house price index.

In addition, temporal aggregation should be considered in constructing housing index. The Malaysian house price index (MHPI) is constructed in every quarter. Therefore, it could lead to the problem of temporal aggregation. This is because temporal aggregation will affect the specification models, estimation of parameters and efficiency of forecasting (Brewer and Kenneth, 1973; Wei, 1979). Furthermore, it may induce bias in risk and return of real estate return and housing index. Geltner (1993) found that temporal aggregation will lead to smooth and erroneously eliminate noise in house price movements. As a result, the house price movements become less risky and the volatility of house price is understated which may not be true. Geltner (1993) also mentioned that temporal aggregated price index will hamper seasonality analysis of house price movement. For instance, the price movements of January, February and March are assumed to be equal and Owusu-Ansah (2013) claimed that this assumption is questionable since the supply and demand of housing market is changing over time.
The arguments toward temporal aggregation problem of house price index are highlighted in some previous studies. Calhoun et al. (1995) found that temporal aggregation smooths price variability across time because spot values are averaged over time with empirical study. In contrast to Englund et al. (1999) and Owusu-Ansah (2013) they claimed that the differences of returns and volatilities among different level of temporal aggregation are small. Moreover, according to Calhoun et al. (1995) temporal aggregation is one of the factors to be considered in constructing housing derivatives. Therefore, it is necessary to investigate the importance of temporal disaggregated price index using Malaysian case study.

Seasonality is one of the factors to be considered for constructing house price index. Schweser (2011) defined seasonality as a pattern that tends to repeat from one period to another. Seasonality will have temporary effect on house price movement. In past, Rosen (1979) mentioned that some factors such as household moving pattern, marriages and family formation rates contribute to housing seasonality. Miller et al. (2012) claimed that seasonality problem will happen if the transaction based on a number of factors common to most households such as school cycles, weather, or holidays. However, seasonality effect cannot be considered as real supply and demand forces. Therefore, the seasonal pattern of house price index should be seasonal adjusted in order to reflect the real performance of housing market.

Miller et al. (2012) has conducted an empirical study on seasonality which covered 138 Core Based Statistical Areas across United States and they found that on average home prices peak in June and hit the lowest value in January. Furthermore, another study carried out by Kaplanski and Levy (2012) stated that house prices are higher on average in the summer by 0.86% to 3.75%. Some past studies also indicate the existence of seasonality in house price (Harris, 1989; Reichert, 1990; Karaganis, 2011). Moreover, price index that impose seasonal pattern will be seasonal adjusted before release to public (Wooldridge, 2006; Greene, 2008). The seasonal pattern gives less attention by Malaysian house price index. Moreover, Malaysia has a different culture from other countries in terms of festival, public holiday, school holiday and weather. These will lead to different bargaining power between seller
and buyer as mentioned by Harding et al. (2003). Thus, it is important to explore the behaviors Malaysian market through seasonality analysis.

To conclude, most of the research highlighted the effectiveness of geospatial data, shrinkage and semiparametric estimators in enhancing hedonic price model. However, the improvement of hedonic price index model gives less attention by past research. Therefore, the research can be further extended to improve hedonic price index model. In addition, it is lack of comprehensive study on the improvement of Malaysian house price index. The temporal aggregation and seasonality problems are not considered in constructing Malaysian house price index. Therefore, it is necessary to establish the methodological enhancement of Malaysian house price index.

### 1.3 Research Questions

The research questions related to the objectives are:

**a) Research question 1:**
How to improve hedonic price index model?

**b) Research question 2:**
How to establish the methodological enhancement of house price index?

### 1.3.1 Research Objectives

1. To evaluate hedonic house price index models.
2. To establish the methodological enhancement of house price index.
1.4 Scope and Limitation of Study

This study scope covered four major residential areas of South Johor. It consists of Pulai, Tebrau, Johor Bahru and Plentong which experienced rapid growth of property development under Iskandar Regional Development. The main aim of this research is to focus on the enhancement of hedonic price index. Therefore, this study will give priority to the regions that can produce large sample of housing transactions. According to the property transaction data from Valuation and Property Services Department (JPPH) these regions has produced large transactions volume of double storey terrace house. The Johor Residential Existing Stock Report 2011-2015 which published by National Property Information Centre (NAPIC) indicated 2-3 storey house placed a largest market share (30%) of residential property. Thus, it shows the importance of 2-3 storey houses in residential market of Johor. The time frame in this research is ranged from year 2006 until 2011 due to the issue of data availability.

1.5 Significance of Study

This section will emphasize the contribution of this study towards the society and knowledge. In term of contribution to the society, house price index plays an important role in delivering the information of housing performance to stakeholders. It helps policy maker to formulate national economic policy respect to housing and property development. Furthermore, house price index could be used to strengthen the consultation report of real estate professionals such as property feasibility report. In addition, house price index is important for investors in term of formulating property investment strategies. House price index will unveil property market behavior in long cycle or short cycle which important for buyer and seller in obtaining a good deal. In future, house price index can be transformed into financial product such as housing derivatives in order to benefit the financial market in Malaysia. However, the benefits of house price index would not be realized without a reliable and house price index. Therefore, it is necessary to revisit the index construction methodology in order to yield a better Malaysian house price index.
This study has contributed knowledge in term of methodological enhancement of hedonic house price index. The weaknesses of existing methodology in constructing hedonic house price index are highlighted in previous section. For example, spatial heterogeneity, multicollinearity, temporal aggregation and seasonality problem has directly affect the reliability of hedonic house price index. Therefore, this study has integrated spatial model, shrinkage estimation, semiparametric estimation, temporal aggregation and seasonality analysis in improving existing house price index. Furthermore, the evaluation method of hedonic price index model in previous research is still unclear. Thus, this study has revisited and established a comprehensive methodology for evaluating the reliability of hedonic price index model. Finally, this study contributed the methodological enhancement of hedonic house price index by using Malaysia case study.

1.6 Research Framework

In order to achieve the two objectives in scientific manner, this section has developed a research framework for this study. The research framework mainly comprised of research issues, research objectives, literature review, research methodology, data analysis, conclusion and recommendation.

Firstly, the research issues and research objectives have highlighted the importance of enhancing hedonic price index. Secondly, by integrating the literature review from various sources it used to determine the elements in improving the hedonic house price index model. Then, the research methodology section will identify the process of obtaining optimal house price index. After that, it followed by data analysis. The result of data analysis will be used to establish the methodological enhancement of house price index. Finally, the research will be ended with conclusion and recommendation. This section will emphasize the knowledge contribution of this study as compared to past studies. At last, it will recommend future research for strengthening the development of hedonic price index model.
The conventional hedonic price index model could expose to econometric problems easily such as omitted variable bias, multicollinearity, and overfitting issue. Moreover, the temporal aggregation and seasonality problems could produce a mislead house price index.

**Objective 1:**
To evaluate hedonic house price index models.

**Objective 2:**
To establish the methodological enhancement of house price index.

**Literature Review**

**House Price Index Construction Methods:**
- Price Index Formula.
- Discussion on price index construction methods.
- Hedonic modellng.
- Hedonic variables.

**Optimal Housing Index:**
- Previous studies of housing index improvement.
- Development of spatial hedonic price index model.
- Incorporation of shrinkage and semiparametric estimator in optimizing hedonic price index model.
- Temporal aggregation and seasonality analysis.

**Research Methodology**

**Data Collection and Management:**
- Historical housing transaction of double storey terrace houses from year 2006-2011.

**Research Process:**
- Hedonic price index development.
- Evaluation methods.
- Temporal aggregation and seasonality analysis.

**Data Analysis**

- Apply simple hedonic model and spatial hedonic model to construct housing index.
- Perform simple hedonic model and spatial hedonic model using shrinkage estimator.
- Perform simple hedonic model and spatial hedonic model using semi-parametric regression.
- Identify an optimal hedonic price index model through evaluation methods.
- Establish the methodological enhancement of house price index.
- Optimize house price index through temporal aggregation and seasonality analysis.
- Adopt the optimal hedonic price index model to construct housing index.

**Conclusion and Recommendation**

Figure 1.1 Study Flow Chart
1.7 Thesis Structure

Chapter One: Introduction

This chapter is initiated by background of study and problem statements in order to emphasize the importance of enhancing the methodology of house price index. After that, it followed by forming research objectives that able to close the research gap. Due to time constraints of the study, the scope and limitation of study are written in this chapter. Moreover, it included the contributions of this study toward the society and the body of knowledge. Finally, a research framework is attached in order to conduct this research in a scientific manner.

Chapter Two: House Price Index Construction Methods

The second chapter has included all essential parts of housing index literature in order to achieve the objectives of the study. The initial part of this chapter will deliver the concept and role of housing index for a country. After that, it followed by the exploration on available index construction methods for housing. It is consists of the strengths, weaknesses and improvements on the existing index construction methods. Besides, it revisited the effectiveness of hedonic regression method for developing Malaysian House Price Index. Meanwhile, the selection of hedonic variables is discussed in this chapter.

Chapter Three: Optimization of Hedonic House Price Index

The third chapter has included literatures of hedonic housing index improvement. There are three major issues to be explored in improving hedonic price index model such as omitted of spatial variables, model complexity and functional form misspecification. Besides, the following section unveiled the methods to identify temporal aggregation and seasonality problem. Both are external factors that can affect the reliability of house price index.
Chapter Four: Research Methodology

Chapter four consists of research methodology that illustrated the process of obtaining optimal house price index. The initial part covered description of study area, data collection, and data management. After that, it followed by descriptive analysis and hedonic price index model development. This chapter has included evaluation methods for measuring the reliability of hedonic house price index model. Finally, the chapter is ended with the computation of temporal aggregation and seasonality analysis.

Chapter Five: Data Analysis

The main aim of this chapter is to explore an optimization methodology for constructing Malaysian house price index. There are two sections in this study. The first section will incorporate geospatial data and two types of estimator such as shrinkage and semiparametric estimation to enhance simple hedonic house price index model. The reliability of various hedonic models is evaluated based on the 1) predictive power, 2) accuracy of price index and 3) level of index revision. In addition, the accuracy of price index is measured by a new method known as double transformation method which had been mentioned in previous chapter. The second section has employed temporal aggregation and seasonality analysis for establishing the methodological enhancement of house price index.

Chapter Six: Conclusion and Recommendation

This chapter has matched the research findings and research objectives in order to give a comprehensive summary and conclusion of the study. Besides, the contributions of this study are included in this chapter in order to emphasize the importance of enhancing hedonic price index model. Finally, the last section of this chapter provides limitations and recommendations for improving this research in the future.
REFERENCES


Jani Venter (2008), The Applications of Commercial Real Estate Derivatives In Investment Strategies, CB Richard Ellis.


