A Study on Spatial Dependence of Housing Prices and Housing Submarkets in Tainan Metropolis, Taiwan

Bor-Ming Hsieh
Department of Land Management and Development,
Chang Jung Christian University, Tainan, Taiwan
bmhsieh@mail.cjcu.edu.tw

Abstract

The purpose of this study is to apply different methods including statistical and spatial analysis techniques to delineate spatial submarkets of housing prices and to examine spatial dependence of housing prices. The data comes from housing transaction prices in the central development areas of greater Tainan City, 2009. Greater Tainan city is a new metropolis amalgamated from former Tainan city and Tainan county. Due to the amalgamation of municipalities, the local government boundaries should be adjusted, and in the mean time, it is worthy to identify spatial submarkets of housing prices in greater Tainan City, compared to former political boundary submarkets. It was found that spatial submarkets of housing prices classified by cluster analysis and spatial techniques are similar. Higher housing prices are concentrated in the core of central development areas while lower prices spread widely around outer ring of the central development areas in greater Tainan city. In testing spatial autocorrelation of housing prices, it was found that it exists significant spatial dependence between housing prices. In modeling housing prices, the results show that spatial submarkets derived by spatial autocorrelation techniques have stronger and higher impacts on housing prices, and the model also have better goodness-of-fit compared to other two types of models.

Keywords: Housing prices, Spatial dependence, Housing submarkets, Cluster analysis, Spatial autocorrelation method.

I. INTRODUCTION

Housing prices are varied by different locations and therefore, can be classified into different spatial submarkets. In many cases, spatial submarkets of housing prices are classified on the basis of physical characteristics of residential dwellings, geographical areas, political boundaries, or market areas as perceived by real estate professionals (Goodman and Thibodeau, 1998; Bourassa et al., 2003). The concept of housing submarket has been extensively applied to housing price research. Housing prices are typically modeled by using hedonic model for estimating individual price data. In a hedonic price model, it is assumed that prices of housing characteristics are similar within a submarket, errors are more likely to be correlated within submarkets than across submarkets. Therefore, controlling for submarkets in hedonic model can substantially reduce estimation errors. The simple way for controlling submarket is to set a series of dummy variables for the submarkets, estimating a separate price equation for each submarket, or adjusting predicted values using errors within each submarket (Bourassa et al., 2007).

However, previous studies have argued that the use of predefined geographical or political boundaries for submarkets in the hedonic price model cannot optimally delineate the impact of spatial attributes on housing prices (Bourassa et al. 2003). As a result, some studies used alternative methods such as Factor Analysis, Principle Component Analysis and Cluster Analysis to define housing submarkets (Dale-Johnson, 1982; Hoesli and Macgregor, 1995; Maclennan and Tu, 1996; Bourassa et al, 2003). Other studies suggest that the use of spatial techniques in hedonic price estimation can significantly reduce spatial dependence of housing prices and have better estimation accuracy, which of these studies provide alternative interpretation of spatial submarkets (Basu and Thibodeau, 1997; Dubin et al., 1999; Case et al., 2004; Bourassa et al., 2007).

The purpose of this study is to apply different methods including statistical and spatial analysis techniques to delineate spatial submarkets of housing prices and to analyze spatial dependence of housing prices. The data is collected from housing transaction prices of Tainan city, Taiwan during 2009. Tainan city, located in southern Taiwan, is one of the oldest cities in Taiwan and is famous in cultural and historical preservation. Now the city is the second largest city in southern Taiwan. In June 2009, the central government permitted the amalgamation of Tainan city and Tainan county into greater Tainan city or called Tainan metropolis. Due to the amalgamation of municipalities, local government boundaries should be adjusted, and in the mean time, it is worthy to identify spatial submarkets of housing prices in greater Tainan City, compared to former political boundary submarkets.

The paper is structured as follow. The next section proceeds with a review of relevant literature on housing submarkets and on spatial analysis of the housing market and the housing price. In Section 3 and 4, respectively, the methods and data are discussed. The empirical results are analyzed in the following section, while the final section provides our conclusions.

I. LITERATURE REVIEW

The review of literature contains two parts. The first part of this section discusses the identification of housing submarkets; the second part discusses spatial analysis of the housing market and the housing price.

1. The Identification of Housing Submarkets

The segmentation of housing market and the identification of submarket has raised a lot of debates over the past several decades (Palm, 1978; Whitehead, 1999; Watkins, 2001). One critical issue is the definition of housing submarkets. Housing submarkets are typically defined on the basis of housing type, the socio-economic characteristics of geographical areas, political boundaries or other real estate professionals (Bourassa et al., 1999). Earlier studies have indicated that the structural characteristics of a dwelling are important in determining housing submarkets (Rapkin et al., 1953; Grigsby, 1963). On the other hand, some studies indicate that the spatial aspect of neighborhood and accessibility attributes of a dwelling are more important than physical structure in determining housing submarkets (Goodman, 1981; Michaels and Smith, 1990). Other studies suggest that housing submarkets are generated by a complicated process which both structural and spatial characteristics of a dwelling should be put into consideration (Adair et al., 1996; Watkins, 2001).

Traditionally, hedonic model has been extensively applied for the interpretation of housing submarkets. Butler (1980) indicates that within housing submarkets, housing prices are similar because submarkets contain close substitutes therefore, implicit prices of housing characteristics are similar for the same reason. Goodman (1981) defines housing submarkets as groups of contiguous local government areas and the author uses hedonic model to analyze the impact of submarkets on housing prices. His results show that the attribute prices of each submarket are not stable across submarkets or time. In addition, Allen et al. (1995) classifies housing submarkets by various dwelling types and the author also uses hedonic model to examine housing price differences between each submarket. The results show that it exists significant distinct rental submarkets based on dwelling types.

Furthermore, some studies have attempted to use statistical methods for delineating housing submarkets. Dale-Johnson (1982) measures the dimension of housing market segmentation by using factor analysis. In his study, five factors are extracted and used to define 10 submarkets. The results show that it exists a significant price difference across submarkets. There are some studies using cluster analysis technique to define housing submarkets. For example, Abraham et al. (1994) uses cluster analysis to analyze metropolitan housing market in the US, while Hoesli et al. (1997) also employs the same method to investigate local real estate markets in the UK. A few studies use composite methods to define housing submarkets. Maclennan and Tu (1996) use principle component analysis to explain the highest proportion of the variation in the data, and then use cluster analysis to identify housing submarkets in Glasgow. It was found that in cluster analysis, the K means algorithm is better than the Ward means. Bourassa et al. (1999) also uses principle component analysis to extract a set of factors from the data and then use cluster analysis to classify housing submarkets in Melbourne, Australia. They also found that some submarkets classified by principal analysis and factor analysis have better results than conventional spatial defined submarkets in house prices estimation.

2. Spatial Analysis of the Housing Market and the Housing Price

In addition to the use of hedonic model and statistical analysis for defining housing submarkets, a variety of studies have attempted to use spatial statistical techniques for analyzing housing markets and housing prices over the past two decades. With respect to spatial structure of the housing market, Adair et al. (1996) indicate that previous studies classified housing markets into various small, well-defined homogenous areas, and these stratified housing submarkets can be applied into wider parts of urban areas but cannot be spatially confined to small neighborhoods. As a result, they suggest that the analysis of housing market should consider a wider spatial area.

Regarding the estimation of house prices, many studies have applied spatial techniques into the traditional hedonic price model since the late 1980’s. For instance, Can (1990) establishes four different housing price models and applied the Moran Test and the Lagrange Multiplier (LM) to examine the spatial residual autocorrelation of house prices. The results show that spatial dependence exists in the error term of house prices and spatial autoregression models have better explanation powers than OLS regression models. Dubin (1992) uses Kriging method to examine spatial autocorrelation in the error term of house prices. It was found that Kriging approach has advantages in estimating house price model by providing useful locational variables.

Furthermore, Can and Megbolugbe (1997) indicate the importance of spatial dependence on the specification of house price function due to spatial spillover effects in the local housing market. As a result, the authors suggest adding spatial variables in the house price model. By investigating housing transaction prices in Miami, US, their results show that spatial hedonic price models have better model goodness-of-fit and higher estimation accuracy than traditional hedonic price models. Pace and Gilley (1997) also have similar results in a comparison of spatial autoregression models with OLS regression models. By using housing data from Boston SMSA, their results show that estimated errors in spatial autoregression models reduced by 44%, compared to OLS models.

Basu and Thibodeau (1998) use OLS and EGLS (Estimated Generalized Least Square) methods to analyze whether if it occurs significant spatial dependence among housing transaction prices in eight districts of Dallas metropolis. Their results show a strong evidence of spatial autocorrelation existed in house prices of eight submarkets. The Kriged EGLS model has higher estimation accuracy than the OLS model in six of eight submarkets. A following study by Dubin et al. (1999) also use different geostatistical methods to analyze spatial autocorrelation occurred in Dallas's housing transaction prices. The authors suggest that with an improvement of GIS techniques, spatial techniques can be smoothly applied into OLS models to examine spatial autocorrelation problems occurred in house prices. They found that the spatial regression model provides better model goodness-of-fit in house price estimation.

Moreover, Bourassa et al. (2007) investigated whether spatial statistical models provide better estimation results than OLS models. Different from spatial techniques, they use neighborhood dummy variables to represent housing submarket conditions. Their results show that when submarket dummy variables are added to spatial models, they perform better estimation of house values than the OLS model. The authors also indicate that hedonic price models with submarket dummy variables are easier to implement than spatial statistical models. However, some studies like Anselin (2002) and Lipscomb (2006) have indicated that if the house price data contains rich location attributes, then previous non-spatial statistical methods can be applied into house price models to improve estimation accuracy of the model. When there is a lack of spatial and location attributes in house price data, spatial statistical methods are adequate to improve spatial dependence existed in adjoining house prices.

In Taiwan, only a few studies have applied spatial statistical methods into house price model since the last decade (Huang, 2004; Ai, 2005; Hsieh and Chu, 2008, Hsieh and Tseng, 2010). With an improvement in GIS techniques, we have seen more opportunities to use spatial statistical approaches in house price models, and there is a need to engage more efforts in this field in Taiwan housing market.

II. METHODOLEDGE

This study analyzes the impacts of housing submarkets on housing prices by employing various methods for defining housing submarkets. In addition to political boundaries, this study uses statistical techniques and spatial techniques to indentify housing submarkets. Regarding statistical techniques, cluster analysis is undertaken to define submarkets. With respect to spatial techniques, Moran's index is employed to test whether if it exists a significant autocorrelation of housing prices, and then Local Indicators of Spatial Association (LISA) is used to classify housing price submarkets. These methods are discussed as follow.

1. Cluster Analysis

Cluster analysis is a technique used for combining housing price observations into groups such that each group is homogeneous or compact with respect to certain characteristics, and each group should be different from other groups with respect to the same characteristics. This study uses two-step cluster analysis to classify housing price submarkets. In the first step, this study uses hierarchical clustering approach to determine the number of clusters. The Ward's method is undertaken to form the clusters. In the second step, K means method is used to compute the cluster. The Euclidean distance is selected as the measure of similarity.

The Ward's method classifies clusters by minimizing total within-cluster sums of squares. The within-cluster sums of squares are known as the error sums of squares, presented in equation (1).

$$D_{ij} = n_i \cdot \left\| \bar{x}_i - \bar{\bar{x}} \right\| + n_j \cdot \left\| \bar{x}_j - \bar{\bar{x}} \right\|^2 \tag{1}$$

Where D_{ij} denotes error sums of squares, n_i, n_j present the number of observation of i cluster and j cluster, respectively. $\bar{x}_i, \bar{x}_j$ denote the centroid of i and j, respectively, while $\bar{\bar{x}}$ presents the centroid of i, j clusters.

The average error sums of squares derived from the Ward's method are then used to be initial cluster centroids in K means method. The euclidean distance is employed to compute the distance of each observation from the centroid, and until none of observations are reassigned and the change in cluster centroids is zero. The euclidean distance is presented in equation (2).

$$D_{ij} = \left(\sum_{k=1}^{p} (X_{ik} - X_{jk})^2 \right)^{1/2} \tag{2}$$

Where D_{ij} is the distance between observation i and j, X_{ij} is the value of the *k*th variable for the *i*th subject, X_{jk} is the value of the *k*th variable for the *j*th subject, and p is the number of variables.

2. Spatial Analysis Methods

The Moran's Index and the Local Indicators of Spatial Association (LISA) are two well-known methods to examine spatial autocorrelation among house prices. If the Moran's index shows that it exists a significant spatial autocorrelation among housing prices, then LISA can be applied to analyze the spatial concentration of higher prices and lower prices. In a spatial context, the spatial concentration of higher or lower housing price can be used to indentify spatial submarkets of housing prices. Their function forms are presented as follows, respectively.

(1) The Moran's Index

The Moran's Index is estimated on the basis of covariance. The function form of the Moran's Index is presented as in equation (3) (Anselin, 1992).

$$I = \frac{N}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \times \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \mu)(x_j - \mu)}{\sum_{i=1}^{n}(x_i - \mu)^2} \text{，for } i \neq j \tag{3}$$

Where the observed variable is the house price; N denotes the sample size; x_i represents the house price in i's spatial unit; x_j represents the other house prices based on i spatial units within certain boundary; μ represents the average house price; w_{ij} denotes locational proximity matrix, and also represents spatial weight coefficients in spatial units.

The value of Moran's Index is between -1 and 1. House prices are positive correlated while the index value is greater than 0; house prices are negative correlated while the index is smaller than 0. The index values are approached to 1 or -1 meaning higher degree of spatial concentration existed in house prices.

(2) Local Indicators of Spatial Association

The method detects whether if it exists a significant spatial dependence of housing prices in a certain boundary, and it can also analyze clusters of spatial concentration of higher or lower housing prices. In particular, the results of LISA can be clustered which are useful to indentify spatial submarkets of housing prices. The LISA statistic can be carried out for a local Moran where the function form is presented in equation (4) (Anselin, 1995).

$$I_i = x_i \sum_j w_{ij} x_j \tag{4}$$

As stated in equation (3), x_i represents the house price in i's spatial unit; x_j represents the other house prices based on i spatial units within certain boundary. w_{ij} denotes locational proximity matrix, and also represents spatial weight coefficients in spatial units.

III. DATA AND VARIABLES

1. The Data

As stated earlier, due to the amalgamation of municipalities, the former boundary of Tainan city should be adjusted to include its adjacent areas. As a result, this study selects 12 districts including six districts of former city districts and another six districts from adjacent areas of former country areas. These 12 districts can be seen as the core development areas of greater Tainan city[1] (see Figure 1). The data comes from housing transaction prices collected by the Ministry of the Interior. The Ministry of the Interior publishes housing transaction price data quarterly in whole Taiwan area since the 1970s. From 1989, the housing transaction price data is free to download from the internet. In 2009, there are 1,385 valid housing transaction price observations collected by this study. The allocation of these observations is presented in Figure 2.

[1] In 2009, about three fourth of housing transaction data is concentrated in these 12 districts, and as a result these districts are selected as the study areas.

Figure 1 The Study Areas

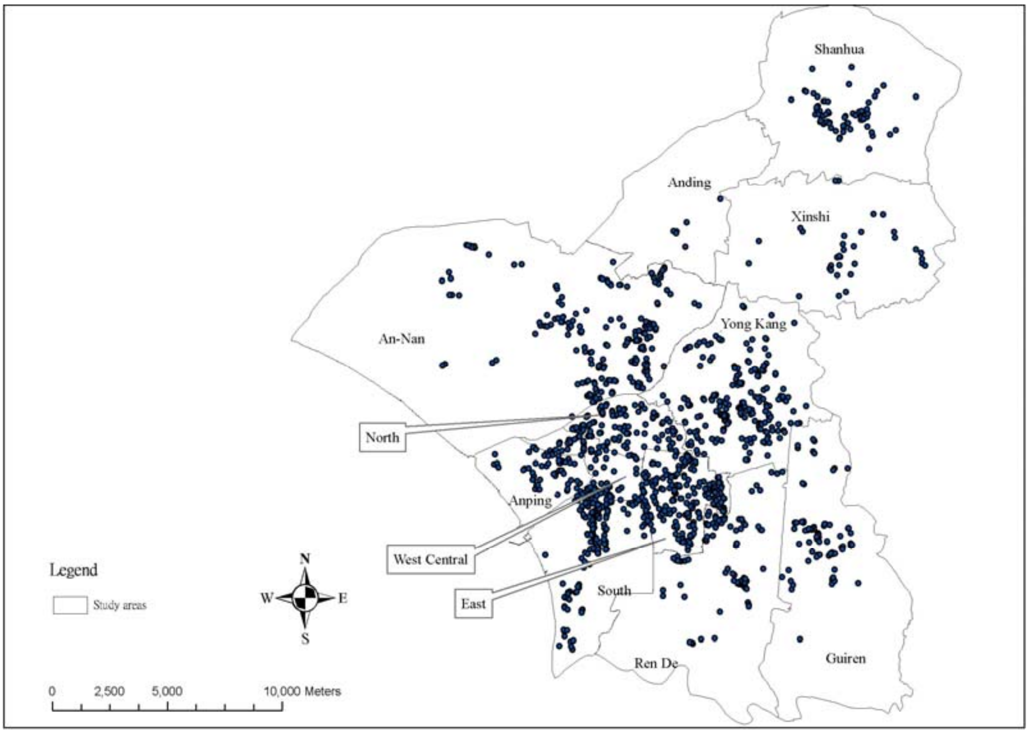


Figure 1. The Allocation of Housing Transaction Price Observations

2. The Variables

The dependent variable is the housing transaction price. Fourteen independent variables are selected in the study presenting physical, neighborhood, and location attributes of housing prices. The descriptive statistics of these variables are presented in Table 1. In the central development areas of Tainan city, the average housing transaction price is 5.22 million NTD (about 180 thousand USD) in 2009. With respect to physical attributes, the average site area is 95.41 square meters, and the average building floor area is 177.05 square meters. The average dwelling age is 14.97 years old. In neighborhood attributes, the average width of road adjacent to sites is 14.83 meter, and the average distance to city center is 5,642.93 meters[2] There are 31 percent of sample dwellings which are adjacent to primary road. About 92 percent of sample dwellings are located in residential zones, while only 3 percent and 5 percent of dwellings are located in commercial zones and other zones, respectively.

Table 1. Descriptive Statistics of Variables

Variable (unit)	Means	S. D.
Housing transaction price (thousand NT)	5,223.08	3,148.72
Site areas (m^2)	95.41	34.68
Building floor areas (m^2)	177.05	73.09
Dwelling age (years)	14.97	13.37
Width of road adjacent to sites (m)	14.83	7.11
Distance to city center (m)	5,642.93	4,133.30
If the site is adjacent to primary road (yes=1)	0.31	0.46
If it is located in residential zone (yes=1)	0.92	0.27
If it is located in commercial zone (yes=1)	0.03	0.16
If it is located in other zones (yes=1)	0.05	0.02
If it is located in E. N. YK. District (yes=1)	0.407	0.49
If it is located in A. C-W. S. District (yes=1)	0.240	0.43
If it is located in R. G. District (yes=1)	0.123	0.33
If it is located in T. S. P. District (yes=1)	0.077	0.27
If it is located in An-Nan District (yes=1)	0.153	0.36

[2] The city center usually can be viewed as the central business district (CBD), this study uses the highest land price area as the city center. In 2009, the highest land price area was located in the Central West district where it is in the inner city area of Tainan city.

Total number of observations=1,385.
* The exchange rate of NTD to USD is about 1,000 NTD to 34.48 USD.

Regarding the location variables, this study uses local government boundaries as spatial submarket to present locational attributes of housing prices. There are twelve political districts in the central development areas of greater Tainan city. This study combines some homogenous districts into five submarkets as: the East, North and Yong Kang district, the Anping, Central-West and South district, the Rende and Gueiren district, the Tainan Science Park district[3] and the An-nan district. The housing transaction price observations are relatively concentrated in the East, North and Yong Kang district and the Anping, Central-West and South district which account for 40.7% and 24% of total observations, respectively. The Tainan Science Park district has the least number of price data accounting for 7.7% of total observations. The number of Housing transaction price data in An-nan district and Rende Gueiren district accounts for 15.3% and 12.3% of total observations, respectively.

IV. EMPIRICAL ANALYSIS

In this section, we first delineate housing spatial submarkets by using three different approaches; then these various spatial submarkets are used to estimate housing prices and also compare the estimation accuracy of three different housing price models.

1. Spatial Housing Submarkets Classification

This study uses three different approaches to classify spatial submarket of housing prices in central development areas of greater Tainan city. The results are discussed as follow.

(1) Spatial Submarkets by Combined Local Government Boundaries

We first use combined local government boundaries for classifying spatial submarkets. The allocation of housing prices by five combined local government boundaries are shown in Figure 1, and the means of housing prices divided by five combined local government boundaries are presented in Table 2.

The average housing transaction prices in the Anping, Central-West and South district are the highest among the five local government boundary submarkets. As stated above, the area is located in the inner city area where it has prosperous

[3] Tainan Science Park is located in the triangle area of Shanhua, Anding and Xinshi districts and as a result, this study uses Tainan Science Park district to present these three districts.

business activities and therefore, housing transaction prices in this area remains higher level in the city. The average housing prices in the East, North and Yong Kang district ranked the second highest among five submarkets which are barely lower than the Anping, Central-West and South area. The area is the most prosperous area and also is education and cultural center. These two areas can be seen as the core development areas of Tainan city. In addition, the average housing transaction prices in the Rende and Guiren area and the Tainan Science Park area are significantly lower than the first two areas, while An-nan area has the lowest housing prices. These areas can be seen as outer ring of central development areas in the city.

Table 2 Descriptive Statistics of Housing Prices by Combined Local Government Boundaries

District	Observations	%	Means of Housing Prices (thousand NT)	S. D.
East, North and Yong Kang	564	40.7	5,715.21	3,342.22
Anping, Central-West and South	333	24.0	5,727.12	3,801.12
Rende and Gueiren	170	12.3	4,624.71	2,354.91
Tainan Science Park	106	7.7	4,490.47	1,941.62
An-nan	212	15.3	3,968.26	1,727.76
Total	1,385	100	5,223.08	3,148.72

(2) Spatial Submarkets by Clustering Analysis

In cluster analysis, we use five variables including site areas, building floor areas, dwelling age, the width of road adjacent to the site, the distance to city center as selected variables to cluster housing transaction prices into separate homogeneous submarkets. These five variables also have significant impacts on housing prices in hedonic price model. Housing prices are classified into four clusters as shown in Figure 2. The descriptive statistics of housing prices in these four clusters are presented in Table 3.

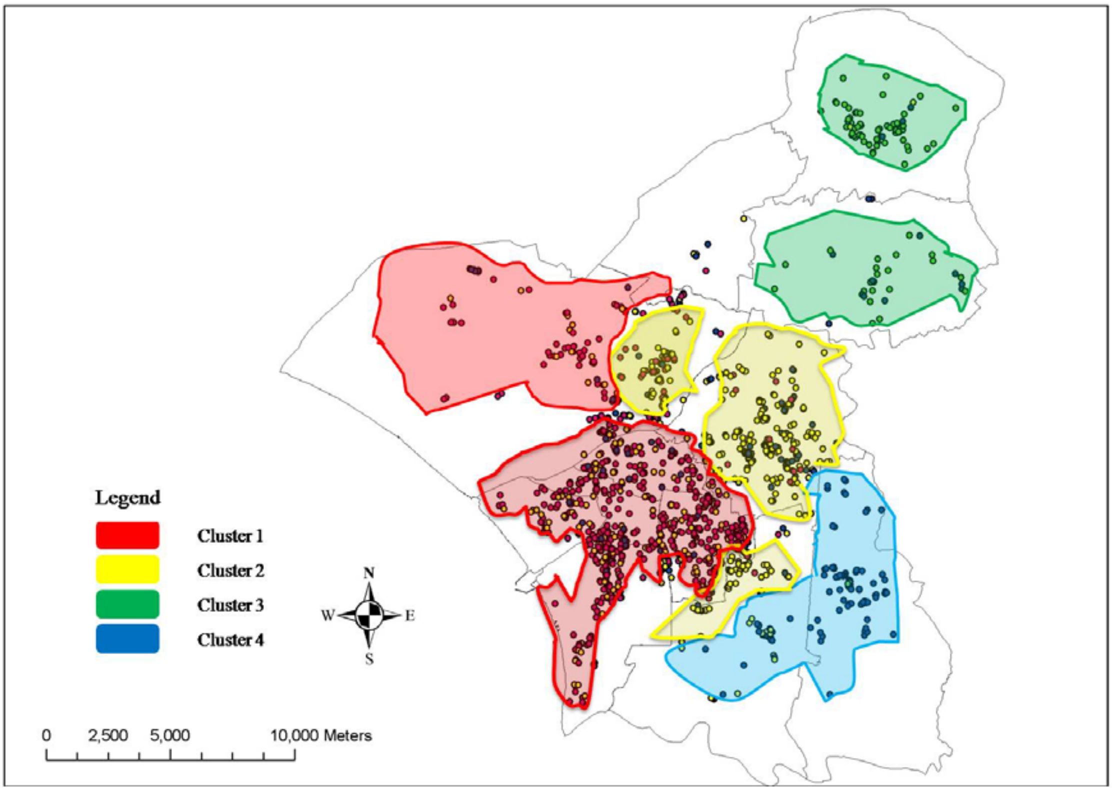


Figure 3 the Clusters of Housing Price Submarkets

Cluster one captures 45.8% of total housing price observations and most of them are located in Central-West, Anping, South, East and also An-nan districts. The average housing transaction prices in this cluster are the highest among four clusters. Cluster two accounts for 32.8% of total observations where they are divided into three zones including Yong Kang, North and East districts respectively. The average housing transaction prices in this cluster are slightly lower than Cluster one. Cluster three is located in the Tainan Science Park area, where it is the outer ring of the central development areas. The average housing transaction prices are the lowest among these four clusters. Cluster four is also located in outer ring areas including Rende and Guiren districts. The average housing transaction prices are slightly higher than Cluster three but are significantly lower the first two clusters. It is clear to note that Cluster one and Cluster two are located in the central areas of Tainan city and have higher housing prices, while Cluster three and Cluster are located in outer ring areas and have lower housing prices. The condition corresponds with Alonso's (1968) rent bid theory in that housing prices are higher in the city center and are declined with an increase in the distance to city center.

Table 3 Descriptive Statistics of Housing Price Clusters

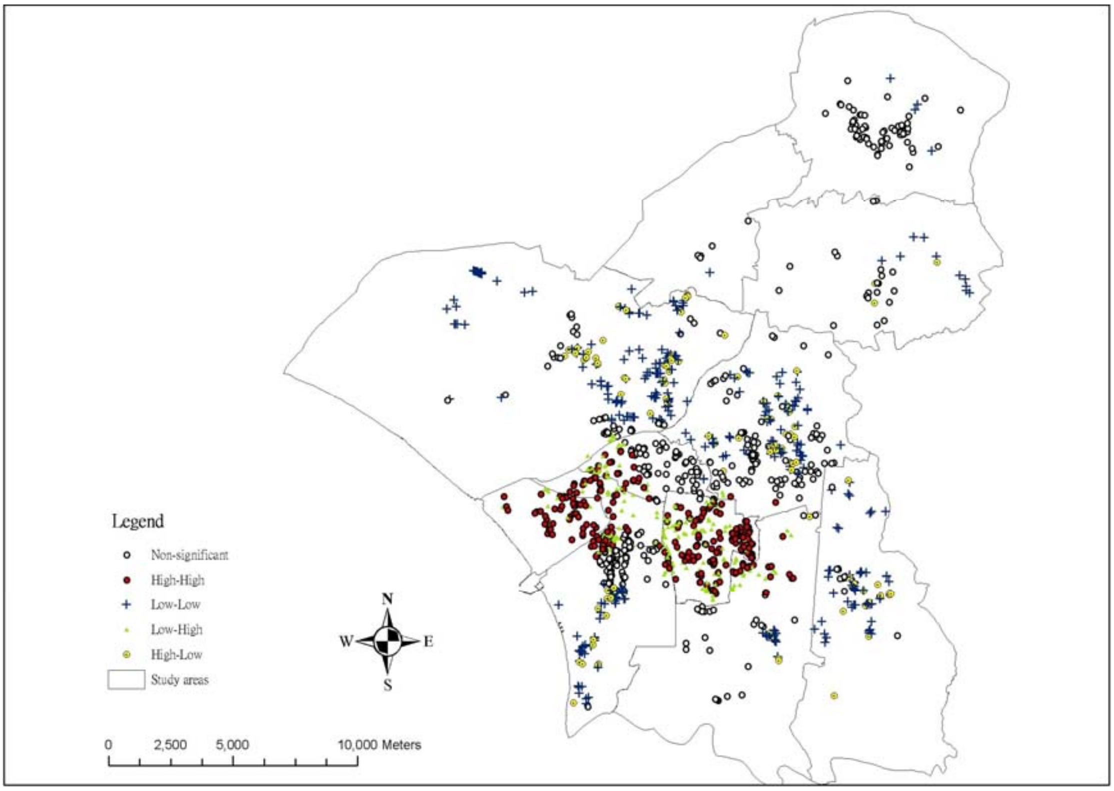
Cluster	Observations	%	Means of Housing price	S. D.
Cluster 1	635	45.8	5,501.17	3,558.94
Cluster 2	454	32.8	5,218.77	2,705.17
Cluster 3	212	15.3	4,623.11	2,971.95
Cluster 4	84	6.1	4,658.45	2,083.70
Total	1,385	100	5,223.08	3,148.72

(3) Spatial Submarket by Spatial Autocorrelation Techniques

This study uses Moran's Index as the global indicator of spatial autocorrelation to examine whether if it exists significant spatial autocorrelation among housing prices. In testing spatial autocorrelation, the first step is to set up a certain boundary. It is assumed that spatial autocorrelation occurred among housing prices within this boundary. Also with an increase in the boundary radius, the degree of spatial autocorrelation of housing prices is reduced. The minimum boundary is set up to satisfy that all housing prices are dependent upon each other. In other words, in the boundary, the number of no adjoining housing prices should be zero. The minimum boundary radius then can be used to examine the local indicator of spatial autocorrelation. The Moran's Index of housing prices in the central development areas of greater Tainan city are listed in Appendix. In 2009, the minimum boundary radius distance that the Moran's Index is set to examine spatial autocorrelation of housing transaction prices is 2,200 meter. It shows significant spatial autocorrelation occurred between each housing price beyond the minimum boundary. Local Indicator of Spatial Association (LISA), is then applied to delineate spatial concentration of housing prices with different levels.

The spatial concentration of different levels of housing prices is presented in Figure 3. It is clear that higher housing prices (High prices surrounded by High prices, H-H) are mainly concentrated in Central-West district, Anping district and East district where they are central business areas as stated above. As presented in Table 4, in higher price zones, average housing prices are about 8.4 million NT (290 thousand USD). On the contrary, lower housing prices (Low prices surrounded by low prices, L-L) are significantly concentrated in Yong Kang district and Anan district. Some lower housing price zones are also located in Rende and Guiren districts and Tainan Science Park areas. These areas are outer ring of the central development areas. The average housing prices in lower price zones are about 3.6 million NT (125 thousand

USD). There exists a significant price gap between higher and lower price zones.



LISA Spatial Concentration of Housing Prices

Table 4 Descriptive Statistics of Housing Prices by LISA

Levels	Observations	%	Means of Housing Prices	S. D.
High-High	284	20.5	8,399.19	4,268.74
Low-Low	339	24.5	3,625.43	1,477.42
Low-High	218	15.7	4,066.10	1,733.05
High-Low	121	8.7	6,074.63	1,480.05
Insignificant	423	30.5	4,723.74	2,528.07
Total	1,385	100	5,223.08	3,148.72

2. Comparison

Three different approaches are employed to identify spatial submarkets of housing prices. The combined local government boundary is a traditional approach to classify administration location of housing prices. Then, the cluster analysis method divides spatial submarket of housing prices by various housing attributes. The spatial statistics approach directly classifies housing prices by using autocorrelation relationship between each housing price. The results of last two approaches regarding statistical and spatial techniques to identify spatial submarkets are similar. Higher housing prices are concentrated in the core of central development areas while lower prices spread widely around outer ring of the central development areas in greater Tainan city. This indicates that housing prices in some areas of central development areas of greater Tainan city should be categorized into the same submarket due to similar location attributes. Moreover, these three types of housing price spatial submarkets are used to estimate their influences on housing prices and also to compare the model's estimation accuracy. The results are shown in Table 5.

Model one uses combined local government boundaries as spatial submarkets to estimate their impacts on housing prices. The R square of this model is 0.703 meaning that the model estimating accuracy is 70.3%. The F value of this model is 271.096 indicating all independent variables together have significant explanatory power on housing prices. All independent variables have significant influences on housing prices. Among these variables, building floor areas have the most significantly positive impacts on housing prices indicating that housing prices are significantly increased with a rise in total floor areas of the dwelling. As expected, an increase in the site area and in width of road have significant effects on a rise in housing prices. Whether if the dwelling adjoins a primary road and whether if the dwelling is located in residential zones or in commercial zones are also have significant positive effects on housing prices respectively. On the contrary, an increase in dwelling age and a rise in the distance to city center have impacts on reducing housing prices. With respect to submarkets, the results show that three of four combined local government boundary submarkets have significant impacts on housing prices. The submarket of the East, North and Yong Kang district has most significant influences on housing prices, following by the Anping, Central-West and South district and the Tainan Science Park district, compared to Annan district.

Model two uses cluster analysis to classify housing prices into four clusters and to estimate their impacts on housing prices. The Model's R square and the F value is 0.687 and 274.219, respectively which is lower than that of model one. This indicates that the estimating accuracy of this model is lower than that of model one. Regarding submarkets, price observations in cluster one and cluster two have significant positive effects on housing prices indicating housing prices in these two clustering areas are higher than other areas. In model three, the model's R square and the F value is 0.784 and 345.1131 respectively which is the highest among these three models. Spatial submarkets classified by spatial autocorrelation techniques have the most important and significant positive effects on housing prices, especially in the higher price zones.

Table 5 A Comparison of Housing Price Models with Three Types of Spatial Submarkets

Model	Model one --using CLGB		Model two --using CA		Model three --using LISA	
Variable	Coeff.	t value	Coeff.	t value	Coeff.	t value
Constant	-2035.11	-5.73***	-2981.81	-5.44***	-965.06	-2.95***
Site areas	54.44	10.47***	44.49	8.61***	50.80	10.60***
Building floor areas	82.32	26.40***	90.93	30.34***	76.65	26.46***
Dwelling age	-36.67	-8.08***	-28.99	-6.52***	-32.53	-7.67***
Width of road	32.70	4.57***	35.95	5.02***	30.27	4.58***
Adjoining to primary road	984.64	8.37***	693.89	6.30***	543.43	5.40***
Distance to city center	-0.088	-4.59***	-0.04	0.89	-0.05	-3.97***
Located in residential zone	595.85	2.42**	802.58	3.39***	534.91	2.47**
Located in commercial zone	1688.34	4.53***	1907.68	5.04***	1178.24	3.40***
Located in E.N.Y. distr.	1410.09	7.39***				
Located in A.CW.S. distr.	1000.56	4.79***				
Located in T.S.P. distr.	1105.91	3.93***				
Located in Anan distr.	230.31	1.03				
Located in Cluster 1			1317.31	3.88***		
Located in Cluster 2			438.75	1.97**		
Located in Cluster 3			-447.47	-1.02		
Located in Higher price zones (H-H)					1701.05	12.40***
Located in Lower price zones (L-L)					-631.28	-5.23***
Located in lower-high price zones (L-H)					424.63	2.93***
Located in Higher-low price zones (H-L)					-204.22	-1.17
R-square	0.703		0.687		0.735	
Adj-R-square	0.701		0.685		0.732	
F Value	271.096***		274.219***		316.572***	
Observations	1,385		1,385		1,385	

In comparison, model three has the highest model estimating accuracy and these spatial submarkets derived by LISA also have most important and significant impacts on housing prices, especially in higher price zones and lower price zones, respectively. This indicates that the use of spatial autocorrelation techniques such as the Moran's

Index and LISA can not only explore spatial dependence of housing prices but also classify spatial submarkets which have better effect in estimating housing prices. In particular, the spatial statistical techniques can clearly identify spatial concentration of higher and lower price zones, respectively. Model one has the second highest estimating accuracy, and these local government boundary submarkets also have important and significant effects on housing prices. This indicates that the use of a set of dummy local government boundary variables to be spatial submarkets also have substantial effects on housing prices. The results also confirm Bourassa et al. (2007) outcomes. The submarkets clustered by various housing characteristics have less effect on housing prices, compared to other two methods. This is probably because this study only uses five housing characteristics to classify housing prices. More housing characteristics would be included or alternative methods should be employed to identify housing prices in further research.

VI CONCLUSIONS

Housing prices are varied by different locations and therefore, can be classified into different spatial submarkets. In many cases, spatial submarkets of housing prices are divided on the basis of geographical or political boundaries. However, the use of these types submarkets would not adequately presents location attributes of housing prices, especially in metropolitan areas. This study employs the cluster analysis approach and the spatial statistical techniques to classify spatial submarkets of housing prices and also makes a comparison to local government boundary submarkets. Our results show that spatial submarkets of housing prices classified by cluster analysis and spatial techniques are similar. Higher housing prices are concentrated in the core of central development areas while lower prices spread widely around outer ring of the central development areas in greater Tainan city. This indicates that housing prices in some areas of central development areas of greater Tainan city should be categorized into the same submarket due to similar location attributes. Furthermore, in testing spatial autocorrelation of housing prices, it was found that it exists significant spatial dependence between housing prices. In modeling housing prices, the results show that spatial submarkets derived by spatial autocorrelation techniques have stronger and higher impacts on housing prices, and the model also have better goodness-of-fit compared to other two types of models. As a result, the spatial technique would be appropriate approach to classify spatial submarkets of housing prices especially in metropolitan areas.

REFERENCES

Abraham, J., Goetzmann, W. and Watcher, S. (1994), "Homogeneous groupings of metropolitan housing markets", *Journal of Housing Economics*, Vol. 3 No. 3, pp. 186-206.

Adair, A., Berry, J. and McGreal, W. (1996), "Hedonic modelling, housing submarkets and residential valuation", *Journal of Property Research*, Vol. 13, pp. 67-83.

Ai, S.L. (2005), "The spatial analysis of land price in residential areas: The example of counties and township areas", Master Thesis, Shih Hsin University. (in Chinese)

Allen, M., Springer, T. and Waller, N. (1995), "Implicit pricing across residential submarkets", *Journal of Real Estate Finance and Economics*, Vol. 11, No. 2, pp. 137-151.

Anselin, L. (1992), "Spatial data analysis with GIS: An introduction to application in the social sciences", working paper, Santa Barbara: National Center for Geographic Information and Analysis.

Anselin, L. (1995), "Local indicators of spatial association—LISA", *Geographical Analysis*, Vol. 27 No. 2, pp. 93-115.

Anselin, L. (2002), "Under the hood issues in the specification and interpretation of spatial regression models", *Agricultural Economics*, Vol. 27, pp. 247-267.

Bourassa, S., Cantoni, E. and Hoesli, M. (2007), "Spatial dependence, housing submarkets, and house price prediction", *Journal of Real Estate Finance and Economics*, Vol. 35, pp. 143-160.

Bourassa, S., Hamelink, F., Hoesli, M. and MacGregor, B. (1999), "Defining housing submarkets", *Journal of Housing Economics*, Vol. 8, pp. 160-183.

Bourassa, S., Hoesli, M. and Peng, V. (2003), "Do housing submarkets really matter?" *Journal of Housing Economics*, Vol. 12, pp. 12-28.

Basu, S. and Thibodeau, G. (1998), "Analysis of spatial autocorrelation in house prices", *Journal of Real Estate Finance and Economics,* Vol. 17 No. 1, pp. 61-85.

Butler, R. (1980), "Cross-sectional variation in the hedonic relationship for urban housing markets", *Journal of Regional Science,* Vol. 20 No. 4, pp. 439-453.

Can, A. (1990), "The measurement of neighborhood dynamics in urban house price", *Economic Geography*, Vol. 66 No. 3, pp. 254-272.

Can, A. and Megbolugbe, I. (1997), "Spatial dependence and house price index construction", *Journal of Real Estate Finance and Economics*, Vol. 14, pp. 230-222.

Case, B., Clapp, J., Dubin, R. and Rodriguez, M. (2004), "Modeling spatial and temporal housing price patterns: A comparison of four models", *Journal of Real Estate Finance and Economics*, Vol. 29 No. 2, pp. 167-191.

Dale-Johnson, D. (1982), "An alternative approach to housing market segmentation using hedonic price data", *Journal of Urban Economics*, Vol. 11, pp. 311-332.

Dubin, R. (1992), "Spatial autocorrelation and neighborhood quality," *Regional Science and Urban Economics*, Vol. 22, pp. 433-452.

Dubin, R., Pace, R. and Thibodeau, T. (1999), "Spatial autoregression techniques for real estate data", *Journal of Real Estate Literature*, Vol. 7, pp. 79-95.

Goodman, A. (1981), "Housing submarkets within urban areas: Definitions and evidence", *Journal of Regional Science*, Vol. 21, pp. 175-185.

Goodman, A. and Thibodeau, T. (1998), "Housing market segmentation", *Journal of Housing Economics,* Vol. 7, pp. 121-143.

Grigsby, W. (1963), *Housing Markets and Public Policy*, Philadelphia, PA: University of Pennsylvania Press.

Hoesli, M. and Macgregor, B. (1995), "The classification of local property markets in the UK using Cluster Analysis", in *The Cutting Edge: Proceedings of the RICS Property Research Conference 1995* Vol. 1, London: Royal Institution of Chartered Surveyors.

Hoesli, M., Lizieri, C. and Macgregor, B. (1997), "The spatial dimensions of the investment performance of UK commercial property", *Urban Studies*, Vol. 34 No. 9, pp. 1475-1494.

Hsieh, B.M. and Chu, S.C. (2008), "A spatial autocorrelation analysis of court auction detached house prices in Tainan City", Conference Proceeding, 2008 Chinese Society of Housing Studies Annual Conference, Tainan. (in Chinese)

Hsieh, B.M. and Tzeng, G.S. (2010), "The changes of spatial allocation of new housing development in Tainan City", 15th Asian Real Estate Society Annual Conference paper, Kaohsiung, Taiwan July 2010.

Huang, S.T. (2004), "The spatial autoregression analysis of housing price in East District of Tainan City", Master Thesis, National Cheng Kung University. (in Chinese)

Lipscomb, C. (2006), "An alternative spatial hedonic estimation approach", *Journal of Housing Research*, Vol. 15 No. 2, pp. 143-160.

Maclennan, D. and Tu, Y. (1996), "Economic perspectives on the structure of local housing systems", *Housing Studies*, Vol. 11 No.3, pp. 387-406.

Michaels, R. and Smith, V. (1990), "Market segmentation and valuing amenities with hedonic models: the case of hazardous waste sites", *Journal of Urban Economics*, Vol. 28, pp. 223-242.

Pace, R. and Gilley, O. (1997), "Using the spatial configuration of the data to improve estimation", *Journal of Real Estate Finance and Economics*, Vol. 14 No.

3, pp.333-340.

Palm, R. (1978), "Spatial segmentation of the urban housing market", *Economic Geography*, Vol. 54, pp. 210-221.

Rapin, C., Grigsby, W. and Blank, D. (1953), *Housing Market Analysis*, Washington, DC: US Housing and Home Finance Agency.

Watkins, C. (2001), "The definition and identification of housing submarkets", *Environment and Planning A*, Vol. 33, pp. 2235-2253.

Whitehead, C. (1999), "Urban housing markets: Theory and policy", in *Handbook of Regional and Urban Economics*, eds by E. Mills and P. Cheshire, Oxford: Elsevier Science.

APPENDIX

The Moran's Index of Housing Prices

Boundary radius distance (m)	Number of No-Adjoining Sample	Moran's Index	Z(I)
100	0.471689	12.541177	511
200	0.378047	14.388436	211
300	0.306034	14.718104	100
400	0.284826	16.338309	55
500	0.254628	16.992329	35
600	0.235960	17.679231	24
700	0.228034	19.029141	20
800	0.217038	20.02823	14
900	0.207337	20.686336	12
1000	0.201242	21.646707	11
1100	0.192719	21.849072	6
1200	0.180368	21.730702	4
1300	0.173101	22.077339	4
1400	0.164608	22.069684	4
1500	0.154092	21.792865	2
1600	0.148592	22.156916	2
1700	0.143266	22.310364	2
1800	0.136614	22.146441	1
1900	0.132105	23.643573	1
2000	0.128711	24.266421	1
2100	0.122219	24.774416	1
2200	0.118471	25.042514	0
2300	0.114794	25.503793	0
2400	0.110346	25.922842	0
2500	0.108803	26.691022	0