ISSN 1666-0040

住宅學報

JOURNAL OF HOUSING STUDIES

第二十二卷第一期 中華民國一百零二年六月
VOLUME 22 NUMBER 1, JUNE 2013

本刊為TSSCI資料庫收錄期刊

中 華 民 國 住 宅 學 會 出 版
PUBLISHED BY CHINESE SOCIETY OF HOUSING STUDIES

住宅學報　第二十二卷第一期　中華民國一百零二年六月　學術論著　第25頁—54頁
JOURNAL OF HOUSING STUDIES, VOLUME 22 NO. 1, JUNE 2013

學術論著

高速鐵路的時空效應對房屋價格的影響—以高雄市為例

Spatial and Temporal Effects of High-speed Rail on House Prices – The Case of Kaohsiung City

廖四郎* 陳靜宜**

Szu-Lang Liao* , Jing-Yi Chen**

摘　要

本文針對交通運輸對房地產價值的影響，提供一個實證的分析。研究的主要目的在於闡述交通運輸的存在對於周邊房地產價值在時間與空間路徑上的變化，了解房屋與高速鐵路的距離及高速鐵路在完成時效上對於房價的影響。先前的研究已針對交通可及性的空間分布或是時間效應對於房地產及土地價值的影響，時間與空間對於房屋價格的二維影響尚未被有系統性的研究。我們使用高雄市在高速鐵路完成時期前後的房屋交易數據，並結合特徵價格法與spline迴歸分析探究房價隨時間與空間二維度的非線性變化。透過空間插值技術-克利金，有效處理空間相關性問題，合理分析房價之動態轉變。最後，本文總結分析在不同距離與時間條件下，追蹤房屋價格對於高鐵影響的變化。

關鍵詞: 房屋價格、時空效應、spline迴歸、高速鐵路、克利金

ABSTRACT

This paper presents an analysis of empirical studies on the effects of transportation facilities on property values. The primary purpose of this study is to address the effects of the presence of transportation facilities on the value of surrounding properties under time and spatial domains. Both the effect of distance from the High-speed Rail (HSR) and the effect of time relative to the completion of the HSR segment contribute to the non-linearity of house prices. While previous studies of the effects of transport accessibility on property and land values have focused on either cross-sectional spatial or temporal patterns, the joint analysis of these two dimensions has not been systematically investigated. We use home transaction data from a period around the completion of the HSR in Kaohsiung City and combine a standard hedonic price model with a spline regression technique to verify non-linear variations of the effect along the temporal and spatial dimensions. The dynamic changes in the house price are analyzed by spatial interpolation techniques (Kriging), which benefit from explicitly dealing with spatial dependence effects. Finally, we estimate the model for a regular series of distances from the HSR, and for time points both before and after the opening of the rail in order to track price effects across various distances and over time.

Key words: house prices, temporal and spatial effects, spline regression, High-speed Rail, Kriging

(本文於2012年2月4日收稿，2012年11月28審查通過，實際出版日期2013年6月)

* 國立政治大學金融學系教授，聯絡作者
Professor, Department of Money and Banking, National Chengchi University, Taipei, Taiwan.
E-mail: liaosl@nccu.edu.tw

** 國立政治大學金融學系博士生
Ph.D. Student, Department of Money and Banking, National Chengchi University, Taipei, Taiwan.
E-mail: tarachen72@gmail.com

1.Introduction

Taiwan High-speed Rail (HSR) started to serve in early 2007. The total length of Taiwan's HSR line is 345km. With an investment cost of approximately US$15 billion, it creates a new record as the world's most expensive build-operate-transfer (BOT) project (Andersson et al., 2010). The new rail link is mainly designed to promote balanced regional development. Currently, Taiwan HSR has eight stations, from north to south, Taipei, Banciao, Taoyuan, Hsinchu, Taichung, Chiayi, Tainan and Zuoying (Kaohsiung), respectively (Fig.1, Taiwan High Speed Rail Corporation, 2012). It had increased transport accessibility along Taiwan's west coast. It also could provide convenient services and make the travel time shorter. One-way travel time between the Taipei and Kaohsiung urban areas has been reduced from about 4 hours to 1.5 hours. It furthermore may improve the strength of land-use around HSR station and increase the land price. Transportation infrastructure investment has often played key role in the restructuring of urban land use and land price patterns, both in Taiwan and somewhere else.

It is commonly believed that effective transportation facility will drive actives to nearby developments and thus rise neighboring housing prices. The belief was intensively studied and many empirical tests conducted using the hedonic pricing model revealed that there is a positive implicit price for effective transportation facilities. However, some studies have indicated that transportation facility would damage nearby housing prices. Transportation facilities affect property values in two ways: positively, via the increased accessibility that they supply, and negatively, due to the negative externalities associated with being close to them. While transportation facilities clearly enhance neighborhood accessibility, they also negatively affect property values by producing noise and pollution (Chernobai et al., 2011; Smersh & Smith, 2000). The transportation accessibility is a critical component in determining land use and property value is well understood. Therefore, to completely understand a relationship between transportation facilities and property values is very important. To the best of our knowledge, distance and time gradients in the effects of transportation facility are expected to be heterogeneity. The related researches have been conducted to dynamical change in house prices that are affected by a transportation amenity and investigate when the price effect appears.

Hedonic price models attempt to explain variation in house prices using property structural, neighboring and location characteristics. In contrast, most structural characteristics are relatively easy to measure and are typically included in publicly available data. Location characteristics such as distance to central business district (CBD), distance to major transportation arteries and so on, are rarely included in publicly available data. However, identifying relevant neighborhood boundaries within urban areas is difficult. Some obvious neighborhood boundaries include municipal and school district boundaries and geographic features like highways, major roads, rivers, and parks. Base on these feats, disclosure the existence of what may be named "spatial effects" have been confirmed. Spatial effects can be thought of as adjacency effects and spatial heterogeneity, respectively. Spatial heterogeneity implies that parameters vary by location and are not homogeneous throughout the data set. In recent years, because of the rapid development of Geographic Information Systems (GIS), spatial data analysis has received considerable attention and played an important role in real estate research. Although many standard statistic techniques are available in traditional data analysis, they can't be implemented uncritically for spatial data (Calderón, 2009). Many research have indicated that tradition hedonic models, even in the presence of structural, neighboring and location variables, didn't sufficiently deal with spatial dependent or spatial heterogeneity issues (such as Basu & Thibodeau, 1998; Dubin, 1998). But using Kriging which has the advantage of giving explicit treatment to spatial dependence effects, in particular when the use of spatially varying attributes is unreliable (because of censoring, measurement errors), or when additional neighborhood attributes are unavailable (Long et al., 2007).

The main objective of this paper is to illustrate how spatial and temporal heterogeneity can be viewed as hedonic price models with spline regression, which refines the inadequacy of standard hedonic price models in spatial and temporal aspects. Then we illustrate how geostatistical approach (Kriging) can fruitfully settle the problem of spatial dependence in study the effect of nearby amenities on property values. The first procedure of this study is to investigate the spatial and temporal heterogeneity of house price responses to transportation system using the example of the Zuoying Sataion (Taiwan HSR), which opened 2007 in Kaohsiung City. We estimate our models for a regular (equal distance interval) series of distances from the HSR, and time points both before and after the opening of the HSR, permitting us to track price effects across various distances and over time. The case study allows us to focus in depth on the issues of spatial and temporal heterogeneity in house price responses. Our focus on observed price change obviates the theoretical question of equilibrium while imposing no assumptions about the direction or magnitude of actual price effects. The second procedure of this study is to demonstrate an analysis of property values for understanding their dynamical change using a case study of the HSR that probably improves the accessibility. Finally, through the integrated analysis between hedonic price models with spline regression and Kriging, we address the question of the impact of spatial heterogeneity and spatial dependence yielding assessment accuracy.

The remainder of this paper is organized as follows. Section 2 reviews related literature and summarizes the theoretical background. Section 3 introduces the outline of the HSR and the study area, and then describes the details of the house price data. In Section 4, house price models are constructed by employing the existing methods of hedonic price models and spline regression for evaluating the effects of the HSR. In Section 5, house prices are interpolated by a geostatistical approach and the dynamical changes in the house prices are discussed. Section 6 presents empirical results for discussion of the objectives. Section 7 highlights the main conclusions of this paper.

Source: Taiwan High Speed Rail Corporation.
Figure 1. Stations of the Taiwan High-speed Rail Line

2. Literature Review

Over the past decade, there has been a great deal of studies using hedonic models to estimate the effect of nearby amenities on house prices. The classical hedonic price model posits a relationship between house price and characteristic variables such as structural characteristics, neighboring characteristics and location factors. Rosen (1974) led to an easier way of attributing effects on property value to characteristics comprising the properties. The coefficients of the hedonic regression are regarded as the implicit prices of the characteristics. Since then, a large literature, including a vast number of empirical studies, has appeared to check the predictability of urban spatial theory by applying hedonic price models (e.g., Chau & Ng, 1998; Can, 1992; Long et al., 2007).

Numerous studies specifically address the marginal effects of the transportation infrastructure on house prices. A number of these studies discussed the positive or negative effects on property values of the addition of nearby rail stations. Debrezion et al. (2007) said the impact of rail stations on property value differs across property types. In general rail stations are expected to have a higher positive effect on commercial properties compared to residential properties for relatively short distances from the stations. In other word, a positive impact of rail station affect commercial property values at short distances and residential values at slightly longer distances. McMillen & McDonald (2004) employed a large sample of single-family house sales to estimate the impact of the New Rapid Transit Line on house prices. The study presents that the house price gradient with respect to distance to the Transit Line stations became strongly negative after the line opened. Kilpatrick et al. (2007) reveal that proximity to the transit corridor (such as superhighways and tunnels) along without direct access conveys a negative impact on nearby house prices. Smersh & Smith (2000) examine the effects of a change in a new bridge because it entailed the possibility of both positive and negative effects resulting from a single change. These effects are enhanced access on one side of the river and increased congestion on the other side. Chernobai et al. (2011) find non-linear in the effects of distance from a new transportation construction. In their study, distance effects are often found to be negative at very short distance, because of noise and pollution externality. At slightly longer distance, the negative effects are lessened, resulting in a positive effect on house prices. At longer distance, the value of the amenity gradually declines to zero.

Equally, the effects of new transportation facilities are not invariable over time, and the effect of the time variable is also non-linear (Chernobai et al., 2011). Smersh & Smith (2000) establish varied effects that were consistent both before and after the opening of the bridge in their studies. In Mikelbank (2005) study, the price impact is not universally positive; the nature of the price impact depends on how long ago the investment was completed. He shows negative effects one year after accessibility, near nil effects at briefly more than two years after accessibility, then peaking at three years after accessibility, and followed by a gradual decline. He cogitates that the negative effect might be due to congestion resulting from the travel adjustment process. The spatial and temporal heterogeneity of house price responses to new transportation facility are detected, hence, we expect the existence of spatial and temporal heterogeneity.

It is well known that there are many researches in which the hedonic regression model has been applied to real estate assessment. Structural characteristics, neighborhood characteristics, and location (accessibility) are the explanatory variables that can be used. Structural characteristics are the individual characteristics of the house itself (age, floor, lot etc.), which can be affected by the property's location. Neighborhood characteristics and accessibility depend on the location of the property. The spatial autocorrelation or spatial dependence of house price is caused by the characteristics that depend on the location. Generally, house price will be directly related to the price of other adjacent houses. Location is probably the most important variable used to explain house price (Kiel & Zabel, 2008). However, if a hedonic model ignores the effects of location, then the residuals of adjacent properties will be correlated (Bourassa et al., 2007). In the recent literature, there is a methodology that considers the spatial autocorrelation of housing prices, geostatistics or spatial statistical (Tsutsumi & Seya, 2008; Chica-Olmo, 2007; Bourassa et al., 2007; Dubin, 1998; Basu & Thibodeau, 1998). Recognizing that the data from many surveys on real estate issues are spatially correlated, we need to extend the traditional statistical methods that are based on the independent assumptions of the observations to the geostatistical techniques that can incorporate the spatially dependent structure of the data for a better assessment.

Since these distance and time gradients in the effects of transportation infrastructure investment are expected to be non-linear, a mainly accomplished methodology for these purposes is spline regression (Chernobai et al., 2011). By using hedonic price models with spline regression in our study, we survey the non-linear in the effects of both distance and time before and after the opening of main transportation facility, the HSR in Kaohsiung City. Such non-linearity has been verified separately in previous studies but only a few analyzed these effects jointly. Furthermore, we apply Kriging technique which has the advantage of giving explicit treatment of spatial dependence effects to demonstrate the dynamical change of property values due to the possibility of improving accessibility. Tsutsumi & Seya (2008) said a lack of sufficient consideration of both the spatial dependence and spatial heterogeneity can lead to serious mistakes in project evaluation based on the traditional hedonic approach. Thus, we believe that our study would make a contribution to the literature on the effects of time and distance from HSR on house prices.

3. Data

Taiwan High-speed Rail started to serve on 2007. It had increased transport accessibility along the western of the island. It also could provide convenient services and make the travel time shorter. It also may improve the strength of land-use around HSR station and increase the land price. In this study, we will discuss the influence of construction of HSR on metropolitan area housing prices by the hedonic price model with a spline regression technique. The Kaohsiung area is chosen as the study area.

The observations on transaction prices and structural characteristics were obtained from Dept of Land Administration, M. O. I.[1] of the central government. Housing data for a total of 5,416 duplex residence transactions in Nanzih district (A1), Zuoying district (A2), Sanmin district (A3), and Gushan district (A4) from first quarter, 2001 to fourth quarter, 2010. This covers two periods, 6 years before the HSR completion and 4 years after the HSR completion. The geographic area of this study is shown in Fig. 2. The Dept of Land Administration provides information for all housing transactions that occurred during this period and it included the structural characteristics of each property such as transaction price, age of house, size of lot and floor, road width and dwelling zone. The detail description of the property characteristics is provided in Table 1. The descriptive statistics for property characteristics are summarized in Table 2.

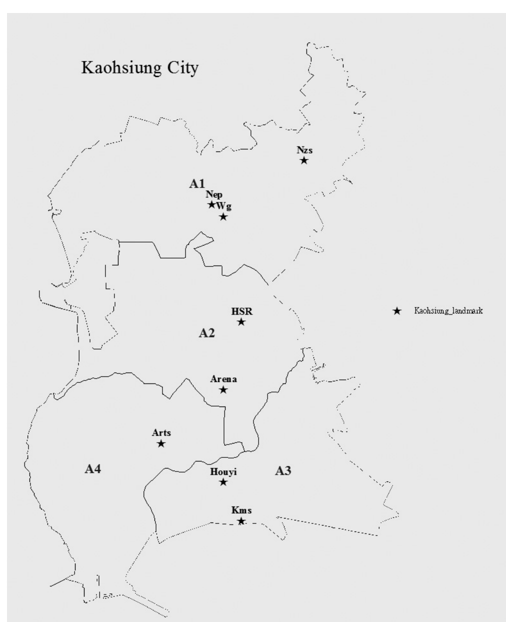


Figure 2. Geographic Area of the Study City

Table 1 Description of Variables

Variable	Description	Units
P	Total house price	10 thousand/NTD
up	House price per square meter	10 thousand/NTD
*P**	The adjusted total house price	10 thousand/NTD
Structural characteristics		
Age	Age of the house	years
Lot	Total area of the lot	square meter
Floor	Total area of the house	square meter
Neighboring characteristics		
Rwidth	Width of the road	meter
Dzone	*Dzone*=1 if the dwelling is in residential zone *Dzone*=0 if the dwelling is in commercial zone	Yes=1; No=0
Accessibility characteristics		
Wg	Euclidean distance to World Games station	km
Arena	Euclidean distance to Hanshin arena shopping plaza	km
Dist	Euclidean distance to HSR station	km
D_1	D_1=1 if the distance to HSR station more than 1.5 km	Yes=1; No=0
D_2	D_2=1 if the distance to HSR station more than 3 km	Yes=1; No=0
D_3	D_3=1 if the distance to HSR station more than 4.5 km	Yes=1; No=0
D_4	D_4=1 if the distance to HSR station more than 6 km	Yes=1; No=0
D_5	D_5=1 if the distance to HSR station more than 7.5 km	Yes=1; No=0
d_2	$D_1\,(Dist - 1.5)$	km
d_3	$D_2\,(Dist - 3)$	km
d_4	$D_3\,(Dist - 4.5)$	km
d_5	$D_4\,(Dist - 6)$	km
d_6	$D_5\,(Dist - 7.5)$	km
Time variables		
Y02, Y03,..., Y10	Year dummies. E.g.,Y02 covers 2002/Q1-2002/Q4	Yes=1; No=0
Time	Number of years between 2001 and trade date of the house	years
T_1	T_1=1 if number of years more than 1 year	Yes=1; No=0
T_2	T_2=1 if number of years more than 2 year	Yes=1; No=0
T_3	T_3=1 if number of years more than 3 year	Yes=1; No=0
T_4	T_4=1 if number of years more than 4 year	Yes=1; No=0
T_5	T_5=1 if number of years more than 5 year	Yes=1; No=0
T_6	T_6=1 if number of years more than 6 year	Yes=1; No=0
T_7	T_7=1 if number of years more than 7 year	Yes=1; No=0
T_8	T_8=1 if number of years more than 8 year	Yes=1; No=0
T_9	T_9=1 if number of years more than 9 year	Yes=1; No=0
Time02	$T_1\,(Time\ \text{-}1)$	years
Time03	$T_2\,(Time\ \text{-}2)$	years
Time04	$T_3\,(Time\ \text{-}3)$	years
Time05	$T_4\,(Time\ \text{-}4)$	years
Time06	$T_5\,(Time\ \text{-}5)$	years
Time07	$T_6\,(Time\ \text{-}6)$	years
Time08	$T_7\,(Time\ \text{-}7)$	years
Time09	$T_8\,(Time\ \text{-}8)$	years
Time10	$T_9\,(Time\ \text{-}9)$	years
Landmark variables		
HSR	Taiwan High Speed Rail Zuoying Station	
Nep	Nanzih Export Processing Zone Station (KRT)	
Wg	World Games Station (KRT)	
Houyi	Houyi Station (KRT)	
Kms	Kaohsiung Station (TRA)	
Nzs	Nanzih Station (TRA)	
Arena	Hanshin Arena Shopping Plaza	
Arts	Kaohsiung Museum of Fine Arts	

Note: d_2,..., d_6 and *Time02*,..., *Time10* denote spline adjustment variables. KRT: Kaohsiung Rapid Transit. TRA: Taiwan Railways Administration.

Using a Geographic Information System(GIS) address matching procedure, all properties are geo-coded; i.e., a latitude and longitude coordinate is calculated for each house. GIS is then used to calculate Euclidean distances from each house to the HSR, World Games Station and Hanshin Arena Shopping Plaza. The distances are measured in km and are labeled *Dist*, *Wg* and *Arena*, respectively, in our estimations.

Table 2　Descriptive Statistics of Variables

		Age	*Lot*	*Floor*	*Rwidth*	*Dzone*	*Wg*	*Arena*	*Dist*	*Time*	*p*	*up*	*p**
A1	Mean	10.73	86.00	163.47	17.20	0.95	2.57	6.50	4.81	6.05	546.44	3.35	524.83
	Median	2.92	79.00	167.60	12.00	1.00	2.11	6.27	4.53	6.25	530.00	3.30	509.25
	Maximum	42.42	341.00	680.12	80.00	1.00	5.26	9.55	7.62	10.00	4869.00	35.34	4869.00
	Minimum	0.00	32.00	38.38	0.00	0.00	0.38	4.19	2.69	0.25	100.00	0.84	102.04
	Std. Dev.	12.36	30.49	62.52	11.39	0.22	1.27	1.33	1.18	2.72	273.30	1.25	271.31
	Observations	1221	1221	1221	1221	1221	1221	1221	1221	1221	1221	1221	1221
A2	Mean	11.48	98.66	213.39	18.38	0.95	4.16	1.44	1.67	5.56	916.59	4.26	812.63
	Median	3.92	84.00	194.64	17.00	1.00	4.13	1.28	1.61	5.50	756.90	3.99	687.82
	Maximum	47.17	1196.25	1613.27	40.00	1.00	11.50	8.16	8.93	10.00	8400.00	45.80	7009.57
	Minimum	0.17	23.00	20.96	2.00	0.00	0.72	0.09	0.26	0.25	130.00	1.18	90.79
	Std. Dev.	12.92	68.89	118.24	8.28	0.23	1.05	0.90	0.78	2.43	718.45	2.26	638.57
	Observations	925	925	925	925	925	925	925	925	925	925	925	925
A3	Mean	25.62	73.62	155.82	14.25	0.67	7.68	3.48	4.84	5.86	545.54	3.59	528.83
	Median	27.67	68.00	140.19	15.00	1.00	7.76	3.30	4.91	5.50	450.00	3.35	435.88
	Maximum	55.00	667.00	3613.03	60.00	1.00	11.39	8.05	8.82	10.00	5300.00	12.80	5372.42
	Minimum	0.08	19.50	37.46	0.00	0.00	4.22	0.61	1.52	0.25	100.00	1.10	101.37
	Std. Dev.	11.62	31.77	109.72	8.17	0.47	0.95	0.96	0.97	2.46	360.42	1.17	352.37
	Observations	2265	2265	2265	2265	2265	2265	2265	2265	2265	2265	2265	2265
A4	Mean	16.56	84.79	186.88	16.92	0.92	6.75	2.70	4.38	5.51	715.17	3.85	851.67
	Median	13.25	75.00	177.72	15.00	1.00	6.77	2.87	4.61	5.50	614.00	3.67	747.33
	Maximum	66.00	534.00	706.69	60.00	1.00	11.17	6.91	8.77	10.00	3630.00	56.75	4600.19
	Minimum	0.00	29.00	19.56	2.00	0.00	4.91	0.78	2.35	0.25	124.00	0.88	139.15
	Std. Dev.	15.55	39.28	101.72	10.60	0.26	1.25	1.27	1.30	2.54	476.73	2.02	560.12
	Observations	1005	1005	1005	1005	1005	1005	1005	1005	1005	1005	1005	1005
Total Study Area	Mean	18.16	82.76	173.14	16.12	0.83	5.76	3.66	4.21	5.79	640.59	3.70	636.30
	Median	21.25	74.39	162.06	15.00	1.00	6.24	3.23	4.51	5.75	537.50	3.49	530.68
	Maximum	66.00	1196.25	3613.03	80.00	1.00	11.50	9.55	8.93	10.00	8400.00	56.75	7009.57
	Minimum	0.00	19.50	19.56	0.00	0.00	0.38	0.09	0.26	0.25	100.00	0.84	90.79
	Std. Dev.	14.42	42.53	103.33	9.61	0.38	2.38	2.02	1.57	2.54	470.30	1.63	466.69
	Observations	5416	5416	5416	5416	5416	5416	5416	5416	5416	5416	5416	5416

Note: Housing data was used for the period from Q1, 2001 to Q4, 2010. The number of included observations= 5,416. A1 denotes Nanzih district, A2 denotes Zuoying district, A3 denotes Sanmin district, A4 denotes Gushan district. The unit of statistics variables is definite in Table 1.

Price(*p*) is the total transaction price of the house which is used as the dependent variable in the hedonic regression estimations in this study. *Age*, shows the age of house, is one of the negative effects affecting house price and we thus expect to have a negative coefficient for *Age*. *Lot* and

Floor are the total areas of the lot and house, respectively, which measured in square meter. The size of each is expected to positively affect the Price. *Rwidth* is width of the road, and it is expected to contribute to a higher house value since this would indicate more convenient of communication. *Dzone* means the dwelling in residential zone, since our samples include residential zone and commercial zone. Thus we use *Dzone* as a dummy variable, *Dzone*=1 if the dwelling is in residential zone, and *Dzone*=0 if the dwelling is in commercial zone. The dwelling in the commercial zone would have higher house price, and thus it is expected to have a negative coefficient for this independent variable. *Wg* and *Arena* represent the distance to World Games Station and Hanshin Arena Shopping Plaza, respectively, which can estimate the effect of nearby amenities on house prices. However, the nearby amenities may increase or decrease housing prices. Hence we do not assume the direction of price effects.

In order to understand how the independent variables impact price, the multicollinearity is worthy of further investigation. Therefore we test the multicollinearity between these independent variables in this model with the correlations matrix in Table 3. As the results, all absolute values of the correlation coefficient are less than 0.5, thus we exclude the multicollinearity problem.

Table 3 Correlation Matrix of Variables used in Estimations

Correlation Probability	*Age*	*Lot*	*Floor*	*Rwidth*	*Dzone*	*Wg*	*Arena*
Age	1						

Lot	-0.294926	1					
	(0)	-----					
Floor	-0.473392	0.575209	1				
	(0)	(0)	-----				
Rwidth	-0.14106	0.074995	0.105976	1			
	(0)	(0)	(0)	-----			
Dzone	-0.283674	0.102422	0.036694	0.008661	1		
	(0)	(0)	(0.0069)	(0.5239)	-----		
Wg	0.360276	-0.121142	-0.023295	-0.08226	-0.308947	1	
	(0)	(0)	(0.0865)	(0)	(0)	-----	
Arena	-0.123165	-0.05022	-0.102548	0.03752	-0.004054	-0.184964	1
	(0)	(0.0002)	(0)	(0.0058)	(0.7655)	(0)	-----

Note: The values are based on the total samples of 5,416 observations. Numbers in brackets are p-values.

4. Constructing house price models

4.1 Preliminary analysis of house price

In order to understand the variance of house price in the study area, we show the change in the average house prices in each district (Fig. 3). Note that not all the house prices in this study area show monochromatic trend (rise or fall) in the period from 2001 to 2010. Thus, under such conditions, it is difficult to intuitively comprehend the impact of the HSR. In order to remove such trends from these data, we use the average price of surrounding area house.

Follow Tsutsumi & Seya(2008) method, we eliminate the trend in house prices of the study area via the house price data of the surrounding area. First, we divide the whole area H into the study area and the surrounding area and we mark as H_A and H_B, respectively. Our study area consists of four districts, namely Nanzih (A1), Zuoying (A2), Sanmin (A3) and Gushan (A4) districts, respectively. The surrounding area is also contains four districts and the relative designated district is shown in Figure 4.

$$H_A = \bigcup_{k=1}^{4} H_{Ak}, H_B = \bigcup_{k=1}^{4} H_{Bk}, H = H_A \cup H_B \quad (1)$$

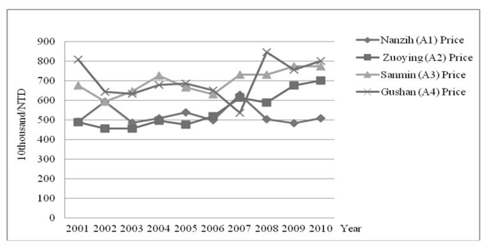


Figure 3. Average House Prices for Each Disctrict by Year

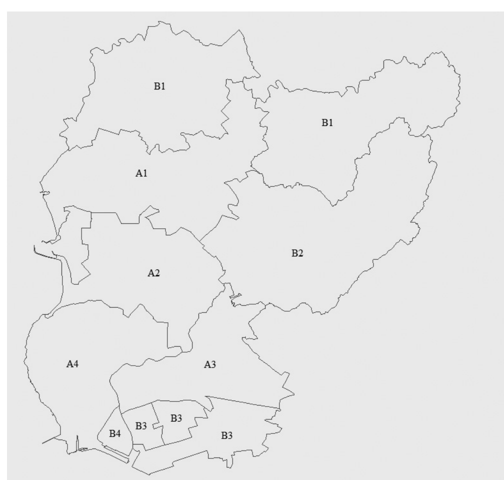


Figure 4. Difinition of the Domains

Subsequently, we can calculate the average house price by the following procedure:

$$\Psi_{H_{Ak}}(t) = \frac{1}{|H_{Ak}|} \sum_{s_i \in H_{Ak}} p_{s_i}(t) \quad (2)$$

$$\Psi_{H_{Bk}}(t) = \frac{1}{|H_{Bk}|} \sum_{s_i \in H_{Bk}} p_{s_i}(t) \quad (3)$$

where s_i expresses the observed sites $\{i = 1,..., 8835(5416+3419)\}$ and t denotes the observed year(t = 2001,..., 2010). $|H_{AK}|$and $|H_{BK}|$ indicate the number of observed sites in H_{AK} and H_{BK}, respectively; $k = 1,...,4$ and p is the house price.

Thus we get the adjusted house price in H_{AK} via the weights as in (4):

$$p^{*}_{s_i \in H_{Ak}(t)} = \frac{p_{s_i \in H_{Ak}(t)}}{\Psi_{H_{Bk}(t)} / \Psi_{H_{Bk}(2001)}} \quad (4)$$

The average adjusted house price in H_{AK} can be calculated as (5):

$$\Psi^{*}_{H_{Ak}(t)} = \frac{1}{|H_{Ak}|} \sum_{s_i \in H_{Ak}} p^{*}_{s_i}(t) \quad (5)$$

4.2. Hedonic price model

We first consider a full hedonic linear regression model over the study area covering the 10-year period:

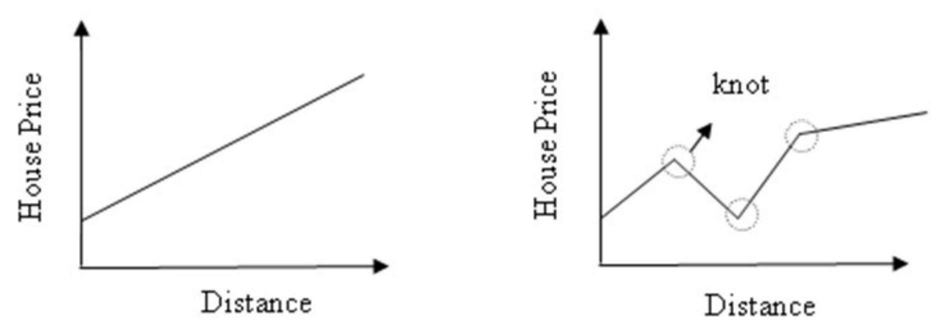
The Full model is shown as:

p = f(Const, Age, Lot, Floor, Rwidth, Dzone, Wg, Arena, Y02, Y03, Y04 Y05, Y06, Y07, Y08, Y09, Y10)* .... (6)

This regression model includes the structural, neighboring and accessibility (excluding *Dist* and spline adjustment variables) characteristics expressed previously. It also contains dummy variables related to the year of transaction that can be viewed as control variables for year of trade.

4.3. Spline regression in hedonic price model

We apply the spline regression technique of Marsh and Cormier (2001) that allows for capturing non-linearity in the effect on a response variable over different segments of the explanatory variable. Spline regression models are used when a regression line is broken into a number of line segments separated by special join points knows as spline knots. At these knots, the slopes of the segments may change, but the end points of the segments are still connected. We give an illustration of two regression models which use diverse methodologies to estimate the effect on the response variable by Fig.5. Hence, spline models are used to require continuity limitations at the join points so that the line can change slope exclusive of unsuitable jumps in the line at those join points.



(a) General Regression Model (b) Spline Regression Model

Figure 5. General Regression Model Versus Spline Regression Model

In order to investigate the effects of HSR and other components on house prices in each of the subsets of the data, we first ignore of the time dummies from the Full regression model and fit the

Section model to the data in each of the subsets in a program. Then we exercise three structures of the Section model. In Section model (a1), no HSR variable is contained. Consequently, it pays no attention to any possible effect of the distance from the HSR on house prices entirely. The Section model (b1) contains all explanatory variables that model (a1) holds and an added *Dist* variable which can identify the linear relationship between the distance to the HSR and the house prices. While transportation facilities improve transportation accessibility, they also negatively affect house prices by traffic congestion. Hence we don't assume the direction of price effects.

The Section model (c1) contains all explanatory variables that model (b1) contains plus extra distance spline variables that can distinguish any non-linear relationship between the different distances to the HSR and the house prices.

In order to regularity between the data subsets, we set the distance cutoffs designed for the spline knots in each subset. The cutoffs are in 1.5 km increments, such as 1.5 km, 3 km, 4.5 km, 6 km, and 7.5 km away from the HSR. In the set of increment aspect, we use a modest increment. Because of the larger increment which hard to estimate the non-linearity in the short distance effects. Similarly, the smaller increment may weaken the statistical results attributable to smaller number of sale transactions inside each distance interval and hence less degrees of freedom in the regression models (Chernobai et al., 2011).

In order to obtain the variation in the effect of the distance to the HSR over time, we divide the dataset into ten subsets - six years before and four years after the HSR opening date. The first six subsets with transaction data from Q1, 2001 to Q4, 2006 cover the construction period of the HSR. The last four subsets are completion periods of the HSR. Then we used hedonic regressions to determine the house value. The value of each element is then determined separately through regression analysis. In our study we apply cross-equation restrictions on parameter values that restrict constant coefficients of the explanatory variables, except the control for the HSR effects across the individual regressions within the program. Only the HSR effects can differ over time. While obtaining a continuous HSR distance effect path in our study is impossible, Chernobai et al.(2011) said that this regression technique becomes a good approximation of the time trail of the HSR distance effect on property values.

By the spline technology, we need to create the spline adjustment variables in Section regression model (c1) in the following way. The five distance cutoffs explained earlier, we first make five dummy variables D_1, D_2, D_3, D_4 and D_5. At this point D_1 =1 when the distance to the HSR is more than the first distance cutoff (1.5 km), and D_1 =0 when the distance is less than 1.5 km. In the same way, D_2 =1 when the distance to the HSR is more than the second distance cutoff (3 km), and D_2 =0 when the distance is less than 3 km, and so on. These dummy variables are then used to create the spline adjustment variables d_2, d_3, d_4, d_5, d_6, corresponding to the second, third, fourth, fifth and sixth 1.5 km distance segments, respectively. At this point $d_2 = D_1(Dist - 1.5)$, $d_3 = D_2(Dist - 3)$, and so on. Thus, for each subset we can obtain Section (c1):

$$P^* = \alpha_0 + \alpha_1 Dist + \alpha_2 d_2 + \alpha_3 d_3 + \alpha_4 d_4 + \alpha_5 d_5 + \alpha_6 d_6 + \alpha_7 X + \varepsilon \quad (7)$$

where X is the vector of the explanatory variables excluding those indicating HSR distance, and ε is the error term. The spline variables can help us to explore the effects on property values at different distances.

In order to determine the total effect of the distance on house prices, for example, if the distance of the observed site to the HSR lies on 3-4.5 km range, we need to summarize the regression coefficients for *Dist*, d_2 and d_3 spline regression variables; views as the total effect of the distance on house prices. It is important to identify that Equation 7 essentially symbolizes six separate equations that correspond to the various distance segments with intercepts $\omega_0, \omega_1, \ldots, \omega_5$ and slopes $\beta_0, \beta_1, \ldots, \beta_5$ as follows:

0-1.5 km range

$$P^* = \alpha_0 + \alpha_1 Dist + \varepsilon = \omega_0 + \beta_0 Dist + \varepsilon \quad (7.1)$$

1.5-3 km range

$$\begin{aligned} P^* &= \alpha_0 + \alpha_1 Dist + \alpha_2 d_2 + \varepsilon \\ &= \alpha_0 + \alpha_1 Dist + \alpha_2 D_1(Dist - 1.5) + \varepsilon \\ &= (\alpha_0 - 1.5\alpha_2) + (\alpha_1 + \alpha_2) Dist + \varepsilon \quad (7.2) \\ &= \omega_1 + \beta_1 Dist + \varepsilon \end{aligned}$$

where D_1 =1 for the distance to HSR station more than 1.5 km.

3-4.5 km range

$$\begin{aligned} P^* &= \alpha_0 + \alpha_1 Dist + \alpha_2 d_2 + \alpha_3 d_3 + \varepsilon \\ &= \alpha_0 + \alpha_1 Dist + \alpha_2 D_1(Dist - 1.5) + \alpha_3 D_2(Dist - 3) + \varepsilon \quad (7.3) \\ &= (\alpha_0 - 1.5\alpha_2 - 3\alpha_3) + (\alpha_1 + \alpha_2 + \alpha_3) Dist + \varepsilon \\ &= \omega_2 + \beta_2 Dist + \varepsilon \end{aligned}$$

where D_1 =1 and D_2 =1 for the distance to HSR station more than 3 km.

⋮

7.5- range

$$\begin{aligned} P^* &= \alpha_0 + \alpha_1 Dist + \alpha_2 d_2 + \alpha_3 d_3 + \alpha_4 d_4 + \alpha_5 d_5 + \alpha_6 d_6 + \varepsilon \\ &= \alpha_0 + \alpha_1 Dist + \alpha_2 D_1(Dist - 1.5) + \alpha_3 D_2(Dist - 3) \quad (7.4) \\ &\quad + \alpha_4 D_3(Dist - 4.5) + \alpha_5 D_4(Dist - 6) + \alpha_6 D_5(Dist - 7.5) + \varepsilon \\ &= (\alpha_0 - 1.5\alpha_2 - 3\alpha_3 - 4.5\alpha_4 - 6\alpha_5 - 7.5\alpha_6) + (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6) Dist + \varepsilon \\ &= \omega_5 + \beta_5 Dist + \varepsilon \end{aligned}$$

where $D_1, \ldots, D_5 = 1$ for the distance to HSR station more than 7.5 km.

Note that the constant term has to be adjusted to adapt the change in the slope. This keeps the regression line continuous even as the regression line pivots change direction as shown in Figure 5(b).

In summary, the three Section regression models on surveying the effects of HSR distance are

Section model (a1): Full model exclusive of the time dummy variable

Section model (b1): Section model (a1) contains the additional variable *Dist*

Section model (c1): Section model (b1) contains the additional distance spline adjustment variables d_2, d_3, d_4, d_5, d_6,

Our first hypothesis is that, if the non-linear effect of the distance on house prices are statistically significant, then we should get a better goodness-of-fit (adjusted r-squared and AIC) better in Section model (c1) than in Section models (a1) and (b1). Our second hypothesis is that, the distance effect would become more obvious for later time periods of the dataset because property holders take time to readjust their house prices in different reaction to HSR.

In order to obtain the varying effects of the time on different HSR distance ranges, we divide the dataset into six subsets. The first range of subset is 0-1.5 km, the second is 1.5-3 km, the third is 3-4.5 km, the fourth is 4.5-6 km, the fifth is 6-7.5 km and above 7.5 km. Similarly, when we investigate the effects of time, we use the same program for each subset. While obtaining a continuous time effect path in our study is impossible, this regression technique becomes a good approximation of the distance trail of the time effect on property values.

The three Section regression models on surveying the effects of time are

Section model (a2): Full model exclusive of the time dummy variable

Section model (b2): Section model (a2) contains the additional variable *Time*

Section model (c2): Section model (b2) contains the additional time spline adjustment variables *Time02,Time03,Time04,Time05,Time06, Time07,Time08,Time09,Time10*

Thus, for each subset we can obtain Section (c2):

$$p^* = \alpha_0 + \alpha_1 Time + \alpha_2 Time02 + \alpha_3 Time03 + \alpha_4 Time04 + \alpha_5 Time05 + \alpha_6 Time06 + \alpha_7 Time07 + \alpha_8 Time08 + \alpha_9 Time09 + \alpha_{10} Time10 + \alpha_{11} X + \varepsilon \quad \text{......................} (8)$$

The variable *Time* shows the number of years from 2001 to the trade date of the house. Additional nine variables, *Time02*, *Time03*, *Time04*, *Time05*, *Time06*, *Time07*, *Time08*, *Time09* and *Time10* are time spline adjustment variables constructed following the spline method explained earlier. At this point, we first make nine dummy variables $T_1, T_2, ..., T_9$. For example, $T_1 = 1$ when number of years more than 1 year, and $T_1 = 0$ when number of years less than 1 year. $T_2 = 1$ when number of years more than 2 years, and $T_2 = 0$ when number of years less than 2 years, and so on. Hence, $Time02 = T_1(Time-1)$, $Time03 = T_2(Time-2)$. For each subset we can obtain Section (c2) as (8). The results of these added regressions are expected to prove whether there existed any differences in the tempo of house value increases over the 10-year period for houses sold in different HSR distance ranges.

5. Exploring the dynamical changes in house price in the study area by Kriging

Firstly, we used geostatistical tools to spatially analyze the Kaohsiung house prices. We first quantified the spatial variations of the observations using the semi-variogram and fitted suitable (in our study we used spherical model) parametric model to the experimental semi-variogram. Then, utilizing the fitted semi-variogram function, we performed ordinary Kriging on the house prices based on the model and produced a set of contour maps.

5.1. Semi-Variogram

Let $p(s_i)$ denote the house price of observed sit s_i . For a stationary process $p(s_i)$, we define the vaiogram, which is also named semi-variogram. The method of estimation for an empirical semi-variogram (Matheron, 1963) is

$$\gamma(\hat{h})=\frac{1}{2NP(h)}\sum_{i=1}^{NP(h)}[p(s_i+h)-p(s_i)]^2= \qquad (9)$$

Where h is the Euclidean distance of the two places (s_i , $s_i + h$), $p(s_i)$, $p(s_j)$ are the house prices at locations s_i and $s_i + h$, respectively. It means that each variable $p(s_i)$ takes different values depending on its spatial location. $NP(h)$ means number of $h(h= \| s_i - s_j \|)$ distant pairs.

In order to perform Kriging for optimal linear spatial prediction, one needs to fit a parametric model over the estimated empirical semi-variogram. Several possible choices of the models including the linear model, exponential model, rational quadratic model, wave model and spherical model are available (Cressie, 1993). In our study, the plot of the empirical semi-variogram $\gamma(\hat{h})$ suggests a spherical model. From the empirical semi-variogram $\gamma(\hat{h})$ for various h, we fitted the spherical parametric model over the empirical $\gamma(\hat{h})$. The spherical model is defined (Cressie 1993) as:

$$\gamma(h)=\begin{cases} 0 & h=0 \\ c_n+c_s\left\{1.5\left(\dfrac{h}{a_s}\right)-0.5\left(\dfrac{h}{a_s}\right)^3\right\}, & 0<\|h\|\le a_s \\ c_n+c_s & \|h\|\ge a_s \end{cases} \qquad (10)$$

Where c_n denotes the nugget effect that measures the micro-scale variation which may be resulted from discontinuous process or measurement-error; $c_n + c_s$ is the sill which represents the variance of the spatial process. If c_s=0, $\gamma(\hat{h})$ is a constant for all $\| h \|$ >0 that indicates no spatial correlation. The range a_s can be interpreted as the house prices are uncorrected beyond distance.

5.2. Kriging

To understand the dynamical changes in house price, we can make interpolated house price maps for the period of interest (Tsutsumi & Seya, 2008). Kriging is a geostatistical technique to interpolate the value of a random function at an unobserved location from observations of its value at nearby locations. It serves to estimate a value at a point of region for which a semi-variogram is known, using data in the neighborhood of the estimation location. Kriging can be classified into simple Kriging, ordinary Kriging and universal Kriging, depending on the kind of random variable processes. Ordinary Kriging is the most suitable estimation method, condition for the random function is second-order or intrinsically stationary, with an unknown mean. In our study we performed ordinary Kriging on the house prices based on the model and produced a set of contour maps, because we use adjusted house price which the random variable presents no drift.

Using Kriging, the house price of the new site s_0 is interpolated by the weighted sum of the observations as

$$P(s_0)=\sum_{i}^{n}\lambda_i P(s_i) \qquad (11)$$

where λ_i is the weight assigned to the indicator at the site s_1.

Kriging technique is an exact interpolation estimator used to find the best linear unbiased estimate. The best linear unbiased estimator must have minimum variance of estimation error.

Mathematically, at location s_0 , we find the linear predictor $\sum_i^n \lambda_i P(s_i)$, where $p(s_1),...,p(s_n)$ are known observations from the survey, that minimizing $E\left[p(s_0)-\sum_i^n \lambda_i p(s_i)\right]^2$, subject to $\sum_{i=1}^n \lambda_i = 1$. Using Lagrange multiplier method, the ordinary Kriging system is

$$\begin{cases} \sum_{j=1}^n \lambda_j \gamma\left[s_i - s_j\right] + \alpha = \gamma\left[s_i - s_0\right], \forall_{i,j} = 1,2,...,n \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad \text{...(12)}$$

In order to achieve unbiased estimators in ordinary Kriging the above set of equations should be solved simultaneously.

5.3. Designing the interpolated house price maps using Kriging

Then, we implement the interpolated house price maps using the above mentioned ordinary Kriging technique. The interpolation was performed by considering the adjusted total house price; however, a map based on the real scale shows to be more helpful for the direct determination of the dynamic impact, therefore, the interpolated results for mapping as in equation (13).

$$\sum_{i=1}^n \lambda_i P^*(s_{i,t}), \qquad (t = 2001,...,2010) \quad \text{... (13)}$$

6. Results

The Full model with year dummies results are summarized in Table 4 by the ordinary least squares regression. About 76% of the variations in house prices are explained by the variations in the selected explanatory variables indicated by the R-squared (adjusted r-squared). The regression coefficients of all year dummy variables are statistical significant at least at the 5% level in 2004, 2007 and 2008. All characteristic variables are significant at the 1% level. As expected, *Age* decreases the house price by about NT$45,175. Both *Lot* and *Floor* positively affect the house price. *Rwidth* statistically significantly increases the price of the house about NT$36,454. The dwelling in the residential has lesser house price than the dwelling in the commercial zone.

The negative statistically significant coefficients for the three out of nine dummy variables used to control the year of sale verify over 10 years house prices. The three dummies *Y04*, *Y07* and *Y08*, have negative coefficients which appear to indicate the house prices of year 2004, 2007 and 2008 are lower than year 2001.

The Full model exclusive of the dummy variables was used to fit the data in the system consisting of the ten subsets. As explained in the earlier section, we include additional variables to control the distance of each house from the HSR. The linear effect of HSR distance on house prices in each subset is measured in a Section regression model (b1), and the non-linear effect in a Section model (c1). Section regression model (a1) does not control the HSR distance. The regression results for the three Section models for each subset are shown in Table 5. The table only presents the results for the adjusted R-squared, AIC (goodness-of-fit) and the HSR distance variable coefficients with the corresponding standard error. Those for the Full model, with and without the HSR distance variables, are also included in the first column of the table for comparison.

Table 4 Hedonic Regression Results for the Full Model

variable	coefficient	std. error
Constant	-20.77799	22.89401
Age	-4.517489	0.278152***
Lot	5.237339	0.090053***
Floor	1.820385	0.041877***
Rwidth	3.64536	0.350722***
Dzone	-42.46692	8.857749***
Wg	19.58313	1.481663***
Arena	-33.88813	1.613368***
Y02	22.48953	20.11712
Y03	-19.27346	19.31562
Y04	-45.12469	16.71217***
Y05	-18.70427	15.64252
Y06	-9.243389	15.39572
Y07	-95.59423	16.2685***
Y08	-36.74509	17.10336**
Y09	2.725155	16.3962
Y10	6.516392	16.46675
R-squared	0.763009	
Adjusted R-squared	0.762306	
AIC	13.69557	

Note: 1. Number of included observations=5,416. The dependent variable is the adjusted total house price p^*. Unit of price is 10 thousand/NTD.

2. ** Denotes 5% statistical significance; *** Denotes 1% statistical significance

Firstly, the results in Table 5 indicate that for the total dataset and for the ten subsets, the adjusted R-squared of the hedonic regression model some increases slightly when house prices are controlled for the distance from the HSR (except for 2008). In each case the value of adjusted r-squared is higher in Section (b1), which includes the linear distance variable, *Dist*, relative to Section (a1). AIC index also shows decline slightly (except for 2002 and 2008). The positive sign of the *Dist* coefficient suggests that the farer distance from the HSR the higher the house price. Based on this system of regressions, we have not seen that HSR improve house prices lying at short distance.

Next, regression model (c1) results indicate that all subsets (except for 2001 and 2002) fit the data better while those control for non-linear effects of the distance using the spline technique. The effect of the spline-line distance from the HSR on house prices varies over the 10-year period contained the construction and completion date can be further understood. While the data does imply that house prices have changed with the HSR construction when factors excluding the distance have been controlled, the local deviation of house price can be extracted.

In 2005 and 2008, the distance spline variable coefficients (*Dist*) are statistically significant for close distances to the HSR. The negative *Dist* coefficient suggests that the house prices decrease as the distance to the HSR increases from 0 to 1.5 km. The degree of this effect becomes largest in 2008, but the effect of distance gets weaker in the following year. In 2009, the distance spline variable coefficient is positive in the 1.5-3 km (d_2). This may be due to a stronger negative effect of the increased congestion externality on selling price of homes lying at close distances from the HSR, thus reversing the effect of distance to the HSR. In particular, the d_2 coefficients from 2003 to 2010 are positive. Then, the d_3 coefficients in 2006, 2009 and 2010 are negative, but the coefficient is positive in 2008. Hanshin Arena Shopping Plaza which lies at south of HSR was open in 2008, we think that this dampens the HSR distance effect. According to the spline methodology, we can generate total effects of the distance on house prices, assuming all else is held constant. For instance, in 2006 houses that lie at 3 km away from HSR are 1.5*323.78=485.66, higher than those next to the HSR. And then, in the 3–4.5 km range the value is 1.5*(323.78-89.29) =351.73. The distance effect appears to be weak for farer distant located house (see Table 6). The data indicates that this distance-related house price appears to move further away from the HSR in 2009. The magnitude of this effect gets larger in the following year. We think that the phenomenon may be caused by the congestion due to the increasing traffic volume over the years, thus raising house prices at the distances further away from the HSR. Table 7 presents annual passenger traffic of HSR in Zuoying station (Ministry of Transportation and Communications R.O.C., 2012). Figure 6 shows the congested roads in our study area (Kaohsiung City Government, 2012). And with Hanshin Arena Shopping Plaza opening, the house prices increase as the distance from the HSR rises.

Table 5 Results of Hedonic Regression Section Models (a1), (b1) and (c1) for Total dataset and Ten Time Subsets

	Total dataset	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Section (a1)		Construction Period						Completion period			
Adjusted R-squared	0.7583	0.3752	0.7941	0.7445	0.8421	0.8344	0.8238	0.5884	0.8830	0.8551	0.8532
AIC	13.7107	14.2402	12.8023	12.9664	13.3364	13.5507	13.5531	14.3863	13.1103	13.1803	12.9900
Section (b1)											
Adjusted R-squared	0.7736	0.3858	0.7943	0.7674	0.8568	0.8429	0.8393	0.6256	0.8828	0.8777	0.8757
AIC	13.6454	14.2265	12.8054	12.8762	13.2403	13.4990	13.4620	14.2933	13.1143	13.0125	12.8250
Dist	**139.94**	**101.87**	27.12	**105.69**	**143.73**	**117.87**	**158.96**	**226.01**	7.05	**170.26**	**169.13**
Std. Error	**(7.30)*****	**(42.11)****	(24.17)	**(20.54)*****	**(19.54)*****	**(17.59)*****	**(16.83)*****	**(28.87)*****	(19.12)	**(15.60)*****	**(15.77)*****
Section (c1)											
Adjusted R-squared	0.7780	0.3819	0.7913	0.7725	0.8637	0.8487	0.8442	0.6285	0.8912	0.8814	0.8789
AIC	13.6265	14.2493	12.8403	12.8718	13.1999	13.4676	13.4365	14.2933	13.0503	12.9890	12.8069
Dist	**-88.17**	4.95	-28.55	-51.02	-101.81	**-134.93**	-79.16	12.35	**-145.28**	-53.97	-91.73
Std. Error	**(25.33)*****	(143.37)	(80.42)	(73.10)	(64.88)	**(64.13)****	(52.07)	(103.35)	**(58.53)****	(58.68)	(64.06)
d_2	**308.01**	144.58	100.61	**227.04**	**362.55**	**315.44**	**323.78**	**242.52**	**213.87**	**302.10**	**352.71**
Std. Error	**(34.19)*****	(208.05)	(110.26)	**(103.11)****	**(89.72)*****	**(85.12)*****	**(70.90)*****	**(133.71)***	**(84.28)****	**(75.79)*****	**(81.80)*****
d_3	**-38.06**	-8.80	-82.65	-53.51	-73.82	-23.35	**-89.29**	45.29	**111.95**	**-91.32**	**-84.97**
Std. Error	**(18.18)****	(113.07)	(61.97)	(54.66)	(47.21)	(42.93)	**(40.89)****	(69.50)	**(52.03)****	**(40.03)****	**(36.09)****
d_4	-4.91	-2.12	43.28	-4.46	43.90	**-55.86**	-38.43	-89.16	-2.49	-20.80	31.48
Std. Error	(13.91)	(86.35)	(46.20)	(45.45)	(35.85)	**(33.41)***	(31.17)	(55.96)	(35.37)	(29.92)	(27.11)
d_5	**70.37**	47.18	-32.91	42.00	96.46	-56.47	50.03	109.51	44.03	**133.57**	37.86
Std. Error	**(25.34)*****	(192.68)	(84.68)	(83.26)	(68.67)	(74.74)	(68.11)	(123.39)	(53.96)	**(46.83)*****	(41.82)
d_6	-64.97	166.54	60.63	181.72	-100.03	**636.66**	-151.70	32.49	-6488.73	**-211.45**	-214.33
Std. Error	(65.71)	(502.57)	(219.70)	(205.93)	(149.56)	**(203.84)*****	(184.74)	(313.49)	(13318.21)	**(111.67)***	(136.42)
N obs.	5,416	290	230	267	524	816	922	616	466	648	637

Note: The dependent variable in each hedonic regression model is (Unit: 10 thousand / NTD). Std. errors are given in brackets. Section Model (a1) contains no HSR distance variables. Section Model (b1) contains linear HSR distance variable, *Dist*. Section Model (c1) contains HSR distance spline variables.* Denotes 10% statistical significance; ** Denotes 5% statistical significance; *** Denotes 1% statistical significance.

Table 6 Non-linear Effects of Distance to the HSR on House Price

Unit: 10 thousand / NTD

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
	---------------------------construction period-----------						-------completion period-------			
0-1.5										
marginal					-134.93			-145.28		
total					-134.93			-145.28		
over 1.5 km					-202.39			-217.91		
1.5-3										
marginal			227.04	362.55	315.44	323.78	242.52	213.87	302.10	352.71
total			227.04	362.55	180.51	323.78	242.52	68.59	302.10	352.71
over 1.5 km			340.56	543.82	270.76	485.66	363.77	102.89	453.15	529.07
3-4.5										
marginal						-89.29		111.95	-91.32	-84.97
total						234.49		180.54	210.78	267.74
over 1.5 km						351.73		270.81	316.17	401.61
4.5-6										
marginal										
total										
over 1.5 km										
6-7.5										
marginal										
total										
over 1.5 km										
7.5-										
marginal										
total										
over 1.5 km										

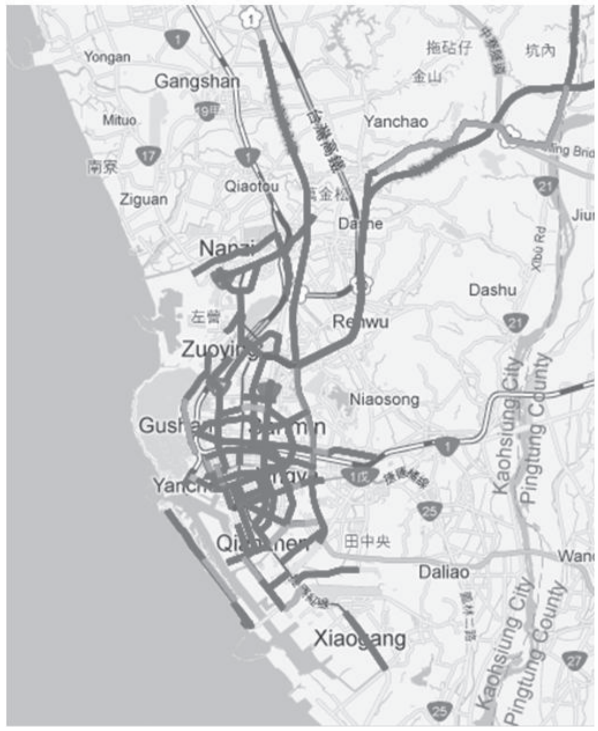
Note: Values are based on the distance spline variable coefficients that are statistically significant at least 10 % level (Table 5). The spline regression parameters indicate how much each distance segment contributed to changing the slope of the regression line. Using these regression parameters, we can calculate the slope coefficient for distance corresponding to each distance segment.

Table 7: Passenger Traffic of High-speed Rail Station (Zuoying Station)

Unit : Passengers

	In	Out
2007	3,226,953	3,358,351
2008	5,815,315	5,930,733
2009	6,137,973	6,152,170
2010	6,523,281	6,524,594

Source: Ministry of Transportation and Communications R.O.C.(2012)



Note: Source: Kaohsiung real-time traffic information. Line color represents different speed km/h.
Freeway: Green line means >80 km/h; Yellow line means 61-80 km/h; Orange line means 41-60 km/h; Red line means <40 km/h.
Artery: Green line means >36 km/h; Orange line means 21-35 km/h; Red line means <20 km/h.

Figure 6. Congested Roads in Study Area

Table 8 presents the results of the hedonic regressions executed in a structure for these six distance subsets and intended to identify the difference in the time path of the house price. The first structure of regressions does not include the time. The second one contains the *Time* variable and determines the linear effect of time, measured in years from the beginning of our sample, on house prices. The third structure is expected to create a more precise time trend as it controls additional time spline variables as the same spline methodology. For instance, spline variable shows the number of years since 2003 until the selling day of a house sold thereafter, and the coefficient indicates the additional yearly transaction price change for houses sold after this date.

For each of the six simultaneous hedonic regression models, the goodness-of-fit in Section (b2) is almost better than the initial model (a2) as is indicated by the values in adjusted R-squared and AIC (except for 3-4.5km). The result is precise when one added includes the time spline variables into the simultaneous regression equations for the entire dataset (except for 7.5- km).

In the Section(b2) that shows a linear relationship between transaction time and house price, the *Time* coefficients indicate statistically significant in house prices over the 10-year period in 1.5-3 km, 4.5-6 km and 6-7.5km distance subsets. Based on the structure the linear effects of time are different, excluding 1.5-3 km, the *Time* coefficients show positive. Based on this system of regressions, we have not seen HSR improve house prices lying at close distances.

However, the structure of the time spline regression equations indicates a non-linear pattern compatible with the one presented in the previous analysis of the effects of the distance on prices along the time dimension. The positive statistically significant time spline variable coefficients are starting in 2005; the positive price was lying within 4.5-6 km. Then, in 2007, the construction was completed; the price was the highest for homes lying within the 0–1.5 km distance range and decreased with distance. This supports the accessibility caused by the new transportation facility. Then, in the 3-4.5km, the time spline variable coefficient is positive performance in 2009 (Table 8). This result is consistent with the prior analysis in distance effect.

Table 8 Results of Hedonic Regression Section Models (a2), (b2) and (c2) for Total Dataset and Six Distance Subsets

		Total dataset	0-1.5	1.5-3	3-4.5	4.5-6	6-7.5	7.5-
Section (a2)	Adjusted R-squared	0.7583	0.9058	0.7892	0.7486	0.7601	0.7058	0.8229
	AIC	13.7107	12.9459	14.1722	13.8305	13.3891	12.3779	13.5211
Section (b2)	Adjusted R-squared	0.7582	0.9060	0.7907	0.7484	0.7605	0.7169	0.8273
	AIC	13.7111	12.9454	14.1663	13.8318	13.3879	12.3415	13.5102
	Time	0.40	4.77	**-11.16**	0.64	**3.82**	**10.92**	-15.81
		(1.29)	(3.24)	**(4.21)*****	(2.64)	**(1.76)****	**(2.47)*****	(10.55)
Section (c2)	Adjusted R-squared	0.7609	0.9237	0.7964	0.7511	0.7621	0.7411	0.82959
	AIC	13.7015	12.7591	14.1490	13.8276	13.3852	12.2702	13.6059
	Time	-14.73	-94.44	129.88	-114.61	-0.74	-83.01	402.90
		(43.37)	(97.38)	(159.42)	(83.24)	(58.43)	(82.34)	(313.34)
	Time02	17.96	113.00	-99.96	102.39	44.19	78.03	-732.76
		(65.30)	(137.52)	(231.56)	(128.80)	(88.03)	(120.65)	(525.65)
	Time03	-41.70	-37.69	-143.28	-50.05	-71.80	-8.85	332.04
		(54.12)	(102.06)	(178.92)	(112.78)	(73.29)	(103.16)	(345.15)
	Time04	24.53	-20.92	95.22	91.74	-2.27	16.60	200.98
		(48.79)	(98.49)	(146.88)	(104.41)	(68.04)	(83.92)	(345.15)
	Time05	**72.06**	99.18	119.18	15.54	**87.05**	18.30	-505.22
		(37.28)*	(88.55)	(110.35)	(79.47)	**(50.45)***	(62.29)	(344.51)
	Time06	**-153.19**	**-267.23**	**-342.33**	-109.63	**-109.89**	-75.68	294.79
		(31.99)***	**(82.74)*****	**(92.20)*****	(69.20)	**(41.72)*****	(59.97)	(294.19)
	Time07	**147.04**	**430.35**	**341.36**	**114.99**	61.90	52.58	291.90
		(32.02)***	**(80.46)*****	**(99.78)*****	**(67.32)***	(41.91)	(53.95)	(289.58)
	Time08	-47.10	-98.59	-31.29	-95.64	4.52	**122.17**	**-550.79**
		(36.53)	(85.43)	(120.67)	(73.98)	(48.94)	**(57.89)****	**(317.84)***
	Time09	27.47	**-253.17**	-106.80	**144.19**	16.62	**-141.29**	321.39
		(36.53)	**(81.42)*****	(119.25)	**(74.00)***	(50.24)	**(53.84)*****	(351.77)
	Time10	-20.41	**150.35**	61.41	**-143.50**	12.30	**98.68**	-46.28
		(36.05)	**(84.09)***	(118.43)	**(72.39)****	(49.54)	**(52.21)***	(327.53)
	N obs.	5,416	390	862	1,430	2,198	480	56

Note: The dependent variable in each hedonic regression model is p^* (Unit: 10 thousand/NTD). Std. errors are given in brackets. Section Model (a2) contains no time variables. Section Model (b2) contains linear time variable, *Time* Section Model (c2) contains time spline variables. * Denotes 10% statistical significance; ** Denotes 5% statistical significance; *** Denotes 1% statistical significance.

In summary, the hedonic regression with spline variables model has showed the nonlinear variation of house price for different distance segments and time lags. We applied spline technique to confirm existence of spatial and temporal heterogeneity phenomenon on prices in our study data. However obtaining a continuous distance or time effect path is not clear enough limited by the regression statistically significant problem. Hence we provide the house price contour map using the above mentioned ordinary Kriging system. The map based on the actual scale shows to be more helpful for the immediate determination of the dynamic impact.

Table 9 shows parameters of semi-variogram in study area. In 2001, the sill variance is the highest, implying the variety in house prices increased. According to the nugget-to-sill ratio which is considered to have strong spatial dependence if the ratio is less than 0.25, and has a moderate spatial dependence if the ratio is between 0.25 and 0.75; otherwise the variable has a weak spatial dependence (Liu et al., 2006). In our study area the house prices have spatial dependence over the 10-year period.

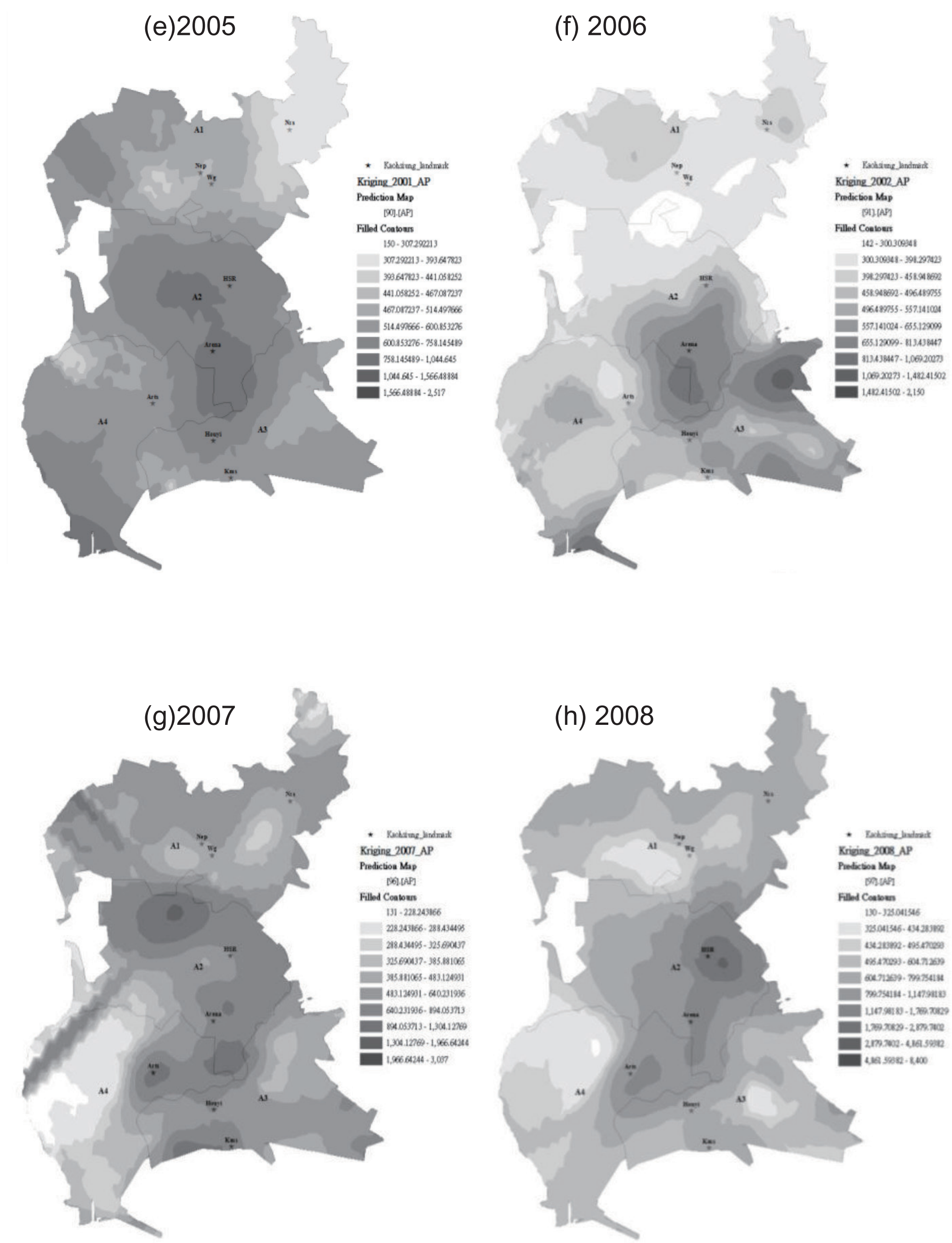
Table 9 Changes in Parameters of Semi-Variogram in Study Area

year	nugget	partial-sill	sill	range	nugget-to-sill ratio
2001	57999.0159	51598.3170	109597.3329	0.0716	0.5292
2002	34175.6205	56196.0314	90371.6518	0.0274	0.3782
2003	0.0000	52228.9948	52228.9948	0.0023	0.0000
2004	45066.1701	305983.0338	351049.2040	0.0025	0.1284
2005	1514.2832	280936.1318	282450.4151	0.0023	0.0054
2006	41844.7989	255334.3573	297179.1562	0.0022	0.1408
2007	107156.6989	100968.5650	208125.2639	0.0202	0.5149
2008	199022.0000	506217.1121	705239.1121	0.0503	0.2822
2009	122877.5792	172937.1284	295814.7075	0.0193	0.4154
2010	53725.9003	120288.8808	174014.7811	0.0041	0.3087

Note: The semi-variogram fitting algorithm performs a least squares fit of various theoretical semi-variograms to an experimental, isotropic semi-variogram. In our study we choose bounded (e.g. spherical) model. The nugget variance means the micro-scale component and a measurement-error component. The sill variance (nugget + partial-sill) means the variance of the spatial process. The range parameter means the extent of the existing spatial correlation.

The Kriging is used for further investigation of the variation of house price in the study area over ten years. Taking into account the interpolated values only, the contour maps of house prices were drawn and shown in Figure 7. As seen from Fig. 7(b), in 2002 the house price is lower in A1 district of the study area, which contains Nep, Wg and Nzs landmarks, and increases rapidly towards the area where among Arts, Arena and Houyi. This indicates that the highest house prices occur at this area prior to the HSR completion. Along with the time course of the HSR completion, house prices increase as the distance to HSR decreases. The results present that the spatial distribution of the house price start to change from 2005, then in 2007, higher house price near to the HSR gradually, as can be seen from Fig. 7(e) and 7(g). Then, from Fig. 7(h), in 2008, the construction was completed; the house prices are higher in the vicinity of the HSR, which indicates that the HSR rises neighboring housing prices. Since the macroeconomic trend of the data is detached, the relative increase in house price is considered to be the result of the HSR operation; the results show that the impact is the greatest around HSR Station. Then, the distance effect appears to weaken for more distant located house. The data indicates that this distance-related house price appears to move further away from the HSR since 2009.

Figure 7. Contour Map of House Price in Study Area



Note: House price (unit: 10 thousand/NTD) has been calculated in order to produce a contour map. Contour map is a map created joining all the points having similar measure. Starting from the spatial distribution of the house price, it has been possible to observe the relationship between the price of observations and their location. The asterisk indicates the location of Kaohsiung facility landmark (include:Nzs, Nep, Wg, HSR, Arena, Arts, Houyi, Kms).

Figure 7. Contour Map of House Price in Study Area (Continuous)

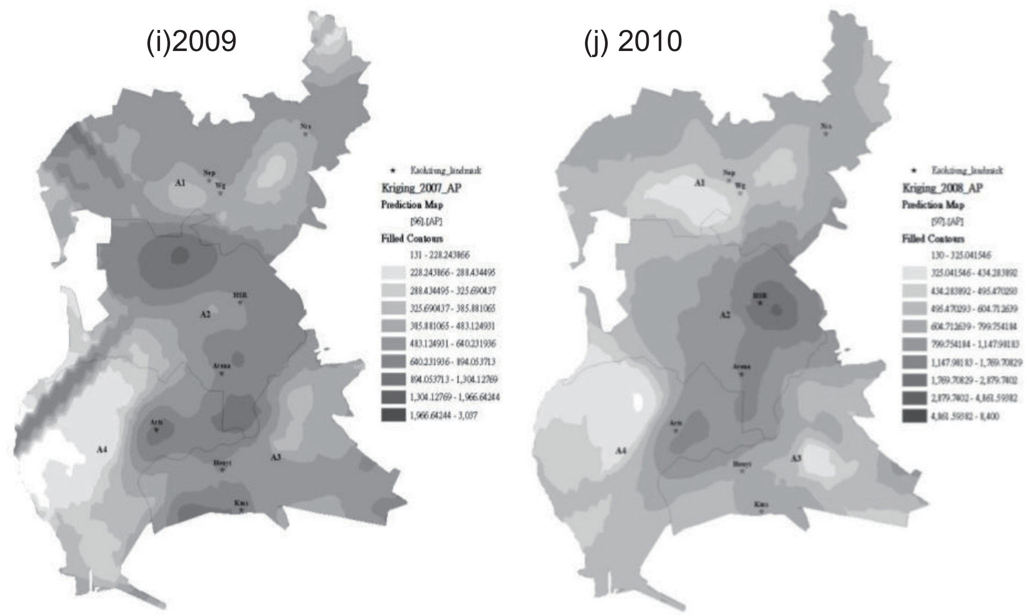


Figure 7. Contour Map of House Price in Study Area (Continuous)

7. Conclusion

The major aim of this paper is to illustrate how spatial and temporal heterogeneity can be viewed as hedonic price models with spline regression, which refines the inadequacy of standard hedonic price models in spatial and temporal aspects. Such issues have been verified separately in previous studies but only a few jointly analyzed completely. Through incremental approach, we first modeled a standard hedonic regression model includes the structural characteristics and dummy variables related to the years of transaction that can be viewed as control variables for year of trade.

Second, in order to obtain the variation in the effect of the distance to the HSR over time, we divided the dataset into ten subsets - six years before and four years after the HSR opening date. We applied cross-equation restrictions on parameter values that restrict constant coefficients of the explanatory variables, except the control for the HSR distance effects across the individual regressions within the program. Then we exercised three structures of the Section model to survey the effect of the HSR distance. In Section model (*a1*), no HSR distance variable is contained which pays no attention to any possible effect of the HSR distance; in Section model (*b1*), with the additional variable *Dist* which can identify the linear effect of HSR distance on house prices; in Section model (*c1*), extra distance spline variables are included, which can distinguish any non-linear relationship between the different HSR distances and the house prices. The linear distance model -Section (*b1*) shows the rise with increasing distance from the HSR for the ten subsets (except for 2002 and 2008). In Section (*c1*), results indicate that all subsets (except for 2001 and 2002) fit the data better when those control for non-linear effects of the distance using the spline technique. The results show, in 2008, the construction completion period, the distance effect is statistically significant in the 0-1.5 km range, but the effect of distance gets weaker in the following year. In

2009, the distance spline variable coefficient is positive in the 1.5-3 km. This may be due to a stronger negative effect of the congestion on selling price of homes lying at close distances from the HSR, thus reversing the effect of distance to the HSR. The data indicates that this distance-related house price appears to move further away from the HSR in 2009.

Third, we transpose the relationship of time and distance in our models. We investigate the effects of time by using the same program for each subset. Section (*a2*) doesn't include the time; Section (*b2*) contains the *Time* variable and determines the linear effect of time; Section (*c2*) is expected to create a more precise time trend as it controls additional time spline variables. Section (*b2*) reflects the linear effects of time are different in six distance subsets, in 1.5-3 km, the *Time* coefficients show negative; in4.5-6 km and 6-7 km, the Time coefficients show positive. Section (*c2*) indicates the positive statistically significant time spline variable coefficients are starting in 2005; the positive price was lying within 4.5-6 km. Then, in 2007, the construction was completed; the price was the highest for homes lying within the 0-1.5 km distance range. This supports the accessibility caused by the new transportation facility. Then, in the 3-4.5km, the time spline variable coefficient is positive performance in 2009. This result is consistent with the prior analysis in distance effect.

Finally, the Kriging is used for further investigation of the variation of house price in the study area over the ten years. Through contour map of house price, we find the spatial distribution of the house price start to change from 2005, then in 2007, higher house price near to the HSR gradually. Then, in 2008, the construction was completed; the house prices are higher in the vicinity of the HSR, which indicates that the HSR rise neighboring housing prices. Since the macroeconomic trend of the data is detached, the relative increase in house price is considered to be the result of the HSR operation; the results show that the impact is the greatest around HSR Station. Then, the distance effect appears to weaken for more distant located house. The data indicates that this distance-related house price appears to move further away from the HSR in 2009.

Based on this research, we believe that this study would make a contribution to the literature on the effects of time and distance from amenities on house prices. Through integrated analysis between hedonic price models with spline regression and Kriging, we address the question of the impact of spatial heterogeneity and spatial dependence yielding assessment accuracy.

Endnote

1: The Department of Land Administration, Ministry of the Interior provides land surveys and registrations, land value assessments, equalization of land rights, land entitlement investigations, land consolidation, land expropriation, land utilization, territorial administration, and other land administration affairs.

References

Andersson, D. E., O. F. Shyr & J. Fu
2010 “Does High-speed Rail Accessibility Influence Residential Property Prices? Hedonic Estimates from Southern Taiwan,” *Journal of Transport Geography*. 18: 166-174.

Basu, S. & T. G. Thibodeau
1998 “Analysis of Spatial Autocorrelation in House Prices,” *Journal of Real Estate Finance and Economics*. 17: 61-85.

Bourassa, S. C., E. Cantoni & M. Hoesli
2007 “ Spatial Dependence, Housing Submarkets, and House Price Prediction,” *Journal of Real Estate Finance and Economics*. 35: 143-160.

Can, A.
1992 “Specification and Estimation of Hedonic Housing Price Models,” *Regional Science and Urban Economics*. 22: 453-474.

Calderón, G. F. A.
2009 “Spatial Regression Analysis vs. Kriging Methods for Spatial Estimation,” *International Advances in Economic*.15: 44-58.

Chau, K. W. & F. F. Ng
1998 “The Effect of Improvement in Public Transportation Capacity on Residential Price Gradient in Hong Kong,” *Journal of Property Valuation & Investment*. 16(4): 397-410.

Chernobai, E., M. Reibel & M. Carney
2011 “Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices,” *Journal of Real Estate Finance and Economics*. 42: 348-370.

Chica-Olmo, J.
2007 “Prediction of House Location Price by Multivariate Spatial Method: Co-kriging,” *Journal of Real Estate Research*. 29(1): 91-114.

Cressie, N.
1993 *Statistics for Spatial Data: Revised Edition*. London: John Wiley & Sons.

Debrezion, G., E. Pels & P. Rietveld
2007 “The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis,” *Journal of Real Estate Finance and Economics*. 35: 161-180.

Dubin, R. A.
1998 “Predicting House Prices Using Multiple Listings Data,” *Journal of Real Estate Finance and Economics*. 17: 35-59.

Kaohsiung City Government
2012 “*Kaohsiung Real-time Traffic Information*,” (http://traffic.kctmc.nat.gov.tw/index_en.jsp).

Kiel, K. A. & J. E. Zabel
2008 “Location, Location, Location: The 3L Approach to House Price Determination,” *Journal of Housing Economics*. 17: 175-190.

Kilpatrick, J. A., R. L. Throupe, J. I. Carruthers & A. Krause
2007 “The Impact of Transit Corridors on Residential Property Values,” *Journal of Real Estate Research*. 29: 303-320.

Liu, D., Z. Wang, B. Zhang, K. Song, X. Li & J. Li
2006 “Spatial Distribution of Soil Organic Carbon and Analysis of Related Factors in Croplands of the Black Soil Region, Northeast China,” *Agriculture, Ecosystems and Environment*. 113: 73-81.

Long F., A. Páez & S. Farber
2007 “Spatial Effects in Hedonic Price Estimation: A Case Study in the City of Toronto,” *CSPA Working Paper Series*, McMaster Univeristy.

Marsh, L. C. & D. R. Cormier.
2001 “*Spline Regression Models*,” (Sage University Papers Series on Quantitative Applications in the Social Sciences, Series no. 07-137) Thousand Oaks, CA: Sage.

Matheron, G.
1963 “Principle of Geostatistics,” *Economic Geology*. 58: 1246-1266.

McMillen, D. P. & J. McDonald
2004 “Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983-1999,” *Real Estate Economics*. 32: 463-486.

Mikelbank, B. A.
2005 “Be Careful What You Wish for: The House Price Impact of Investments in Transportation Infrastructure,” *Urban Affairs Review*. 41: 20-46.

Ministry of Transportation and Communications R.O.C.
2012 “*Monthly Statistics of Transportation & Communications*,” (http://www.motc.gov.tw/en/home.jsp?id=609&parentpath=0,154&mcustomize=statistics301.jsp).

Rosen, S.
1974 “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy*. 82: 34-55.

Smersh, G. T. & M. T. Smith
2000 “Accessibility Changes and Urban House Price Appreciation: A Constrained Optimization Approach to Determining Distance Effects,” *Journal of Housing Economics*. 9: 187-196.

Taiwan High Speed Rail Corporation
2012 “*Stations of the Taiwan High-speed Rail Line*,” (http://www.thsrc.com.tw/en/?lc=en).

Tsutsumi, M. & H. Seya
2008 “Measuring the Impact of Large-scale Transportation Projects on Land Price Using Spatial Statistical Models,” *Paper in Regional Science*. 87: 385-401.