

Hedonic Price Modeling of New Residential Property Values in Xi'an City, China

Kai Liu¹ & Toshiaki Ichinose^{2,3}

¹Doctoral Program in Geoenvironmental Science, University of Tsukuba, Tsukuba City, Japan

²Center for Social and Environmental Systems Research, National Institute for Environmental Studies, Tsukuba City, Japan

³Graduate School of Environmental Studies, Nagoya University, Nagoya City, Japan

Correspondence: Kai Liu, Doctoral Program in Geoenvironmental Science, University of Tsukuba, 1-1-1 Tennodai, Tsukuba City, 305-8577 Ibaraki, Japan.

Received: June 23, 2017

Accepted: July 20, 2017

Available online: September 1, 2017

doi:10.11114/ijsss.v5i9.2510

URL: <https://doi.org/10.11114/ijsss.v5i9.2510>

Abstract

This study analyzed new residential property values of Xi'an City in March, 2014. Results show that accessibility indices, such as distance to CBD, have been well capitalized into the residential property values. Particularly, a within-zone housing unit is sold 9.4% more than if it was outside the attendance zone of a Key Primary School, i.e. home buyers have how much willingness-to-pay for the high-quality educational resource. Another corollary is got that the accessibility to subway stations has not significantly been capitalized, only with a low premium reflecting in the real estate market of Xi'an City. Considered that spatial local singularities caused by unobserved variables or estimation bias can be associated with multi-regression errors, this study herein has explored an unconventional viewpoint to residual problem, which combines the regional differences (coming from real world) and the spatial distributions of singularities (feedback from data). Furthermore, whole samples are classified into 5 agglomerations for revealing the underlying reasons about the future trend and variation of real estate market within each region. It is effective to provide scientific basis of decision making for the real estate investors and planners.

Keywords: Hedonic price modeling, new residential property values, residuals

1. Introduction

Residential property value is one of the most important indicators of urban economic development, and it concerns everybody in a city. Especially in recent years, residential problem in major Chinese cities has generated global interest among investors, policy makers, scholars and journalists (Wu, Deng, & Liu, 2014). In Chinese cities, variations of housing price are directly affected by macroeconomic and policy, apparently showing characteristic of spatial disproportion. Cities with location and resource superiority have a comparatively higher and more rapidly growing housing price, and this brings a bigger pressure on the housing market in the area near the quality public resource and local public goods such as Key Primary Schools (see Note 1), main stations, etc. (Hu, Zheng & Wang, 2014). Zheng and Kahn (2007) proved that subway stations, schools, green spaces and other public resources can be able to explain a quarter of the spatial differences about residential property values in Beijing. Chinese scholars also found that property values of residences inside a Key school's attendance zone had a higher and faster growth rate in Shanghai (Wang, Ge, & Zhang, 2010; Shi & Wang, 2014; Liu & Sun, 2015). In addition, transportation infrastructures, high schools, parks and other amenities have also been proved to be the important determinants of housing prices (Kim & Zhang, 2005; Zheng & Kahn, 2007). Certainly, extensive researches about the impact of public rail transit on housing price and land use pattern have been done in Beijing, Shanghai and various other cities, where the appreciation effect is observed and premiums caused by local public goods are quantitatively estimated (Wang, Zhu, & Zhang, 2004; Pan, Ren, & Yang, 2007). Wei, Wei, & Jiang (2014) exposed that development of urban rail transport has a marked impact on real estate industry. In short, properties bought and sold on real estate market are a kind of compound commodities, whose values are usually influenced by the following two key factors: commuter cost and education environment. It is because that these factors can contain the main part of implicit price used to embody property buyers' convenience will and willing-to-pay for these location accessibilities. Moreover, lower commuter cost and quality education environment are

usually accompanied by other key local public goods. According to related studies, hedonic pricing models, which use such properties as proxies, can often be employed to quantify these location advantages. Thus this study selects several explanatory variables and creates a hedonic pricing model for estimating their capitalization effects in residential property values in Xi'an City, China.

Nevertheless, some limitations are existed and remaining issues must be discussed before using hedonic pricing models. This study has explained main reasons (a Key Primary School, distance to CBD, etc.) that control the new residential property values clearly. However, it cannot be said that the most appropriate regression model was found, even though every T-test is statistically significant if stepwise regression method is used. It is because that there are many observable variables that haven't been added into the model. In statistics, a residual (or fitting error) is an observable estimate of the observable statistical error, or an observed value is the difference between the observed value and the estimated function value (Cook & Weisberg, 1982). Unobserved variables are always existed and ineradicable. While the quantity of residuals can reflect the effectiveness of regression fitting, as well as residuals are related to unobserved variables' existence in a sense. By using residuals analysis, reliability of the data and disturbances from unobserved variables can be analyzed. However in most cases, previous studies just only unilaterally used hedonic price modeling and forced on proving their studies by themselves, but ignored or even dodged residuals problem and may not discuss about the unobservable variables, which is a very important breakthrough in solving the task on the local singularities. Perhaps geographers always tend to subjectively verify the results using regression. Residual analysis is widely employed in the field of remote sensing science. Residuals are usually regarded as an important indicator or variable, for instance, evaluating the filtering effect in geodesy (Xu, Gui, & Han, 2008) and achieving the highway accident forecasting (Li, Huang, Liao, Han, & Yuan, 2005) by using gray residual error model. Through reviewing the recent literature, it is easy to realize that geographers did not pay enough attention to residuals and ignored the academic interest in urban economics. Therefore, residuals analysis is performed following the completion of hedonic regression approach. Results do verify that spatial singularities caused by unobserved variables or estimation bias can be related to residuals. Cluster analysis is widely used to make spatial classification in geography. At last, this study has classified 5 housing agglomerations. Focusing on local differences by using cluster analysis, this study aims to prove some hypotheses which is conceived from the previous results of residual analysis. Trends of housing price and real estate market prospects within each region are investigated. In summary, this study applies a series of statistical techniques, i.e. through a way from regression estimation on an integral macroscopic aspect, to residual analysis on an individual microscopic aspect, and then clustering on a regional aspect. Anyway, this study tried to explore a new and possible hedonic approach, unlike other related literatures just stick to parameter estimation. Besides quantizing the estimation of premium for every characteristic, this study has revealed potential spatial singularities and location differences which can be further probed from hedonic regression.

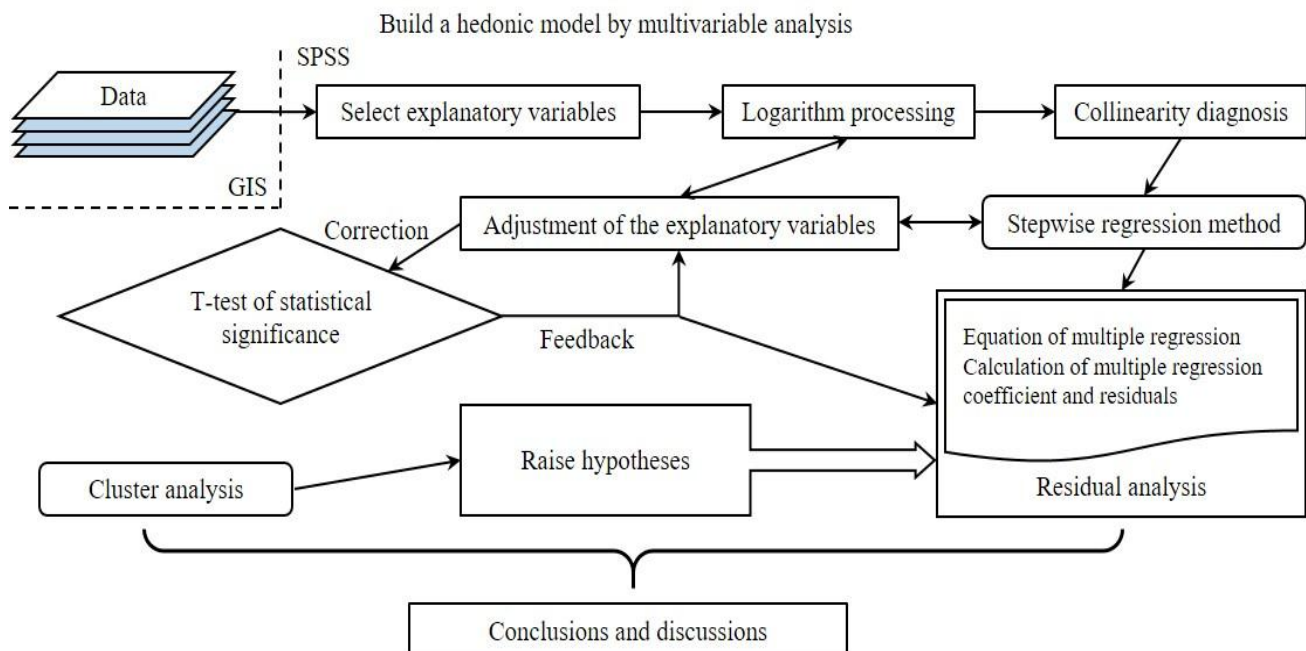


Figure 1. Flowchart of the study

The main objectives of this study are to: 1) attempt to apply hedonic pricing models in housing market of Xi'an City; 2) search for the most suitable regression model by using SPSS and measure the premium by presenting the monetary value of each key characteristic to urban planner and decision-makers for optimizing urban development projects; 3) reveal the causes and mechanism about spatial singularities and differences. 7 chapters are organized. Chapter 1 introduces the research objective and a brief outline of the research. Chapter 2 provides theoretical framework of the hedonic approach. Then, chapter 3 describes the data set, and some pretreatment and explanatory variable selection are done supported by GIS. Additionally, the regression process is discussed in detail. In Chapter 4, conclusions of hedonic pricing model are drawn. Through residual analysis and cluster analysis, remaining issues like spatial singularities and regional differences are solved in Chapters 5 and 6. In particular, some hypotheses are raised Chapter 5 and proved in Chapter 6, respectively. Finally, full summary of the whole article is sorted out in Chapter 7. Figure 1 shows the flowchart of the study and outlines the interrelationships of each step.

2. The Hedonic Pricing Model

2.1 Background of Hedonic Pricing Model

The price of commodity is described by the summation of the price which reflects its characteristics that meet the needs of people, which is the core idea of hedonic pricing model. In fact, researches on hedonic pricing model have been conducted for a long time. Haas (1922a & 1922b) first used this approach to explore the farm sales. Thereafter, the application of hedonic pricing model could be widely expanded in many academic fields such as environmental economics (Ridker & Henning, 1967) and urban economics (Tiebout, 1956; Lancaster, 1966; Muth, 1966). Lancaster (1966) developed what he called a "new theory of consumer demand", in which the then standard microeconomic demand theory was modified by stipulating that what consumers are seeking to acquire is not goods themselves (e.g. cars or train journeys) but the characteristics they contain. Thereby, Lancaster with his new theory of consumer demand can be called the founder in theoretical basis of microeconomics. Since when a concrete form of hedonic pricing model have been put forward by Rosen (1974), it has been frequently applied in many studies. He considered that property is a complicated commodity with many dimensions, and differences in the selling prices of houses will be dictated by a number of factors, including the quality of the housing structure, neighborhood characteristics, accessibility to the central business district (CBD), as well as the environmental (dis)amenities (education, landscape, etc.) associated with the property. Hedonic pricing models use a market good (a residence) as a proxy, within which a non-market good, such as environmental (dis)amenity, is implicitly traded. Public resources like schools, urban green spaces, etc. and accessibility about people's living produce benefits that presumably make a neighborhood a nice place in which to live, and these benefits are reflected in the prices of surrounding properties. Rosen and Fullerton (1977), along with Judd and Watts (1981), found the relationship between property values and schools' teaching quality. They proved that there is a positive correlation at a statistically significant level by using hedonic pricing model. Consequently, the amenity of public resources and other location characteristics could be valued in monetary terms according to how much people pay for such benefits in their housing. That's the reason why premiums are capitalized into a residence.

The development of geographic information systems (GIS) has gradually made hedonic pricing model a powerful tool. Since location characteristics are the most important parameter for real estate valuation, using the framework of GIS can make it possible to calculate accurate, consistent, and unbiased explanatory variables, such as accessibility to subway stations, in a fast and efficient manner (Kong, Yin, & Nakagoshi, 2006). Supported by GIS, researches using hedonic pricing model has experienced unprecedented rapid development in the recent decades.

From the perspective of econometrics, when using hedonic pricing model, several issues such as regression function form, selections of explanatory variables and estimation methods, had been extensively studied by many scholars. For example, here are some empirical researches like house price index construction (Wu et al., 2014), valuation of access to jobs and amenities in housing price (Ding, Zheng, & Guo, 2009), airport noise and residential property values (Nguy, Sun, & Zheng, 2014), effect of urban villages on commercial housing price in time series containing before and after the re-construction policy (Zhang, Zheng, & Sun, 2014), etc.

Since the 1990s, empirical researches on hedonic pricing model have been entering a period of extensive application. Scholars gradually pay more attention to a detailed study of the impact from a certain characteristic on housing price, and this have become one of the mainstream in the researches using hedonic pricing model (Wen & Jia, 2004). These location characteristics can be many kinds of local public goods such as education, medical treatment, traffic, environmental quality, etc.

As Chapter 1 pointed out, properties bought and sold on real estate market are a kind of compound commodities, whose values are usually influenced by the following two key factors: commuter cost and education environment. Focusing on Chinese special education system where the so-called *Xue Qu Fang* problem is always happened (see Note 2), Xi'an City is selected as study area because of the budding rail transportation and city structure as an ancient capital, where is

a typical case for making quantitative analyses to estimate the capitalization effect of local public goods such as CBD, subway stations and Key Primary Schools.

2.2 Structure of the Hedonic Pricing Model

Before using hedonic pricing model, the form for the hedonic price function must be chosen. Housing is not usually considered as a homogeneous commodity (McLeod, 1984). A house buyer purchases both a dwelling and a set of site characteristics. The price at which a residence is traded is depending on lots of factors: housing structure, location, neighborhood, and environmental characteristics. The household must find the dwelling with the best combination of features at the most apposite price (O'Sullivan, 2002). The implicit prices of spillover benefits (premium) from the factors can be inferred based on the hedonic hypothesis (Rosen, 1974). Consequently, the traditional hedonic pricing model takes the following form:

$$P = f(L, S, N) \quad (1)$$

where P is market price of the housing. Here we defined total price of a new residence as the dependent variable P . L , S , N denotes independent variables, which represents location characteristics, structure characteristics, and neighborhood characteristics, respectively (Ridker & Henning, 1967). In this study, 11 variables were chosen as hedonic price indices. Data with detailed descriptions and descriptive statistics, can be found in Table 1.

Table 1. Selection of variables used to be added into the hedonic pricing mode and results of descriptive statistics*

Dependent variable	Unit	Description	Mean	SD	Min.	Max.	Expected sign**
P	10,000RMB	The total price of a new residence	92.25	109.92	17.70	1230.00	
Independent variable							
DC	km	A measure of the ease of accessibility to CBD, i.e. straight-line distance to central business district	8.58	4.88	1.06	32.89	-
DSX	km	A measure of the ease of accessibility to railway station, i.e. straight-line distance to Xi'an Station	8.75	4.73	1.30	30.58	-
DS	km	A measure of the impact of the nearest subway station which is a straight-line distance from the residence	1.37	1.60	0.10	15.73	-
DR	km	A measure of the integrated impact of the nearest river or lake from the residence to the boundary of the nearest river or lake	1.63	1.15	0.01	6.63	-
S	m ²	Gross Floor Area	110.86	48.66	24.02	456.00	+
BR	individual	Number of bedroom	2.47	0.79	0.00	5.00	+
LR	individual	Number of living room	1.93	0.36	1.00	5.00	+
WR	individual	Number of washroom	1.43	0.64	1.00	4.00	+
HB	Dummy variable	Whether the residence is in a high-rise building above 12 floors or not	0.67	0.47	0.00	1.00	-
SN	Dummy variable	Whether the residence is a south facing house or not	0.86	0.35	0.00	1.00	+
E	Dummy variable	A measure of education environment. $E_d=1$, where d is the radius of buffer made by a Key Primary School when the residence is located into this area; or $E_d=0$, when the residence is located outside this buffer.	0.61	0.49	0.00	1.00	+

* Data source from biggest Chinese homepage of real estate website named *An Ju Ke*. (<https://xa.anjuke.com>)

** + and - represent increasing and decreasing effects on the housing price, respectively.

The hedonic pricing model employed in the evaluation of resources typically uses multiple regression techniques to relate property price details to diverse characteristics of differing properties, and sort out the different contributions. Many economists have stressed that economic theory does not suggest an appropriate functional form for hedonic price equations (Rosen, 1974; Freeman, 1979; Halvorsen & Pollakowaski 1981; Cassel & Mendelsohn 1985). Consequently, it is necessary to try reasonable functional forms to each variable and utilize a multiple regression equation (Cassel & Mendelsohn 1985). Halvorsen and Pollakowaski (1981) recommend the use of the Box-Cox flexible functional form in

hedonic analysis, which would measure best performance with goodness-of-fit. This form is commonly applied in the literatures using the hedonic pricing model (Shultz & King 2001, Kong et al., 2007). However, Cassel and Mendelsohn (1985) had pointed out a disadvantage of this methodology: that using a best-fit criterion to choose functional forms does not necessarily lead to more accurate estimates of characteristic prices. Maurer et al. (2004) thought that simplicity of the model and explanatory in economics are more important comparing to the accurate Box-Cox conversion. In general, this study just considers coefficient of determination (R^2) as the indicator to judge whether the function form of hedonic pricing model has measured better performance with goodness-of-fit.

Equation 1 is a very classical defined formula which is frequently referenced. However, as Rosen (1974) pointed out, there is no reason to expect the relationship between the price and factors to be liner. Another situation is that different factors based on different units and orders of magnitude. In most cases, non-liner is to be expected, because purchasers cannot treat individual housing attributes as discrete items from which they can pick and mix until the desired combination of characteristics ID is found (Morancho, 2003). Thus following previous studies (Harrison & Rubinfeld 1978; Can, 1992; Zheng & Kahn, 2013), this study employs a function form combining with logarithmic and semi-logarithmic form, which can be respectively defined as follows:

$$\ln(P) = a_0 + \sum_{n=1}^i a_i \ln(Z_i) + \varepsilon \tag{2}$$

$$\ln(P) = a_0 + \sum_{n=1}^i a_i Z_i + \varepsilon \tag{3}$$

where Z_i denotes explanatory variables, a_0 denotes a constant, a_i denotes regression coefficients, and ε is the residual error term.

This study applies a form which is combined with Equation 2 &3 to deal with a variety of situations which happened correspondingly for the selected variables as the case may be. In other words, some variables must be regarded as dummy variables (0 or 1). This is because logical variables or category variables cannot make logarithmic transformation (Li, Silva, & Bo, 2011). This detail and regression process will be brought to light in the next Section. In this study, real estate transaction prices were collected for Mar. 2014 in Xi'an City. Based on the hedonic pricing model, this study attempts to estimate the capitalization effect of location characteristic such as CBD, subway stations and Key Primary Schools. At the same time, implicit prices (premium) for characteristics specific to the housing itself and neighborhood characteristics are also estimated.

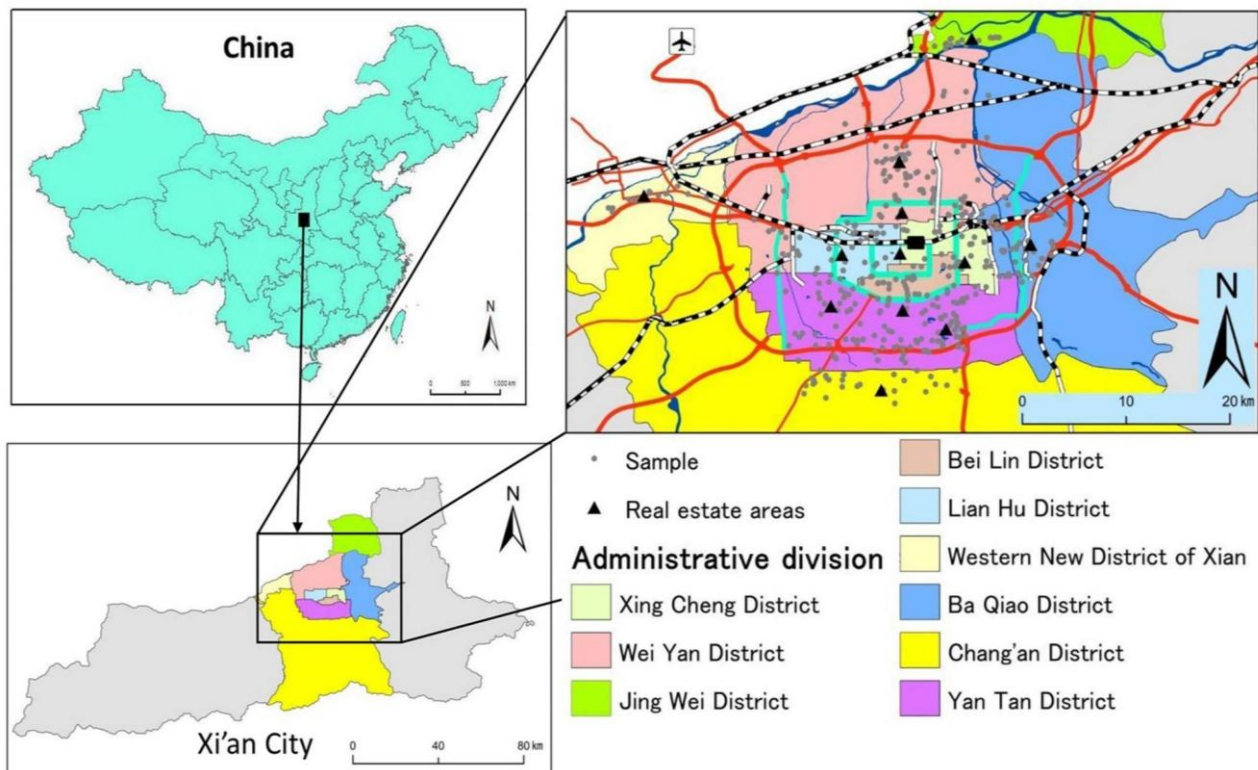


Figure 2. Study area

3. Data Set and Regression Procedure

3.1 Study Area

This study was conducted in Xi'an City, the capital of Shaanxi Province in China (Figure 2). Resident population was 8.47 million in 2010. Total area is 10108 km², among which urban area is 3582 km² in 2012 (Shaanxi Provincial Bureau of Statistics, 2013). Xi'an is located on the northwest of the People's Republic of China, in the center of the *Guan Zhong* Plain. As an ancient capital with rich historical heritage and tourism resources, it has become a leader city of China Western Development Policy. Along with urban development, housing price in Xi'an City has also been grown rapidly in recent years like other Chinese cities. By all accounts, Xi'an City was taken as the study area to confirm the practicality of hedonic pricing model in this study. Furthermore, comparing with other previous studies in China, which usually took eastern three metropolises like Beijing, Shanghai, and Guangzhou et al. as their study areas, this study tries to get an original case by taking Xi'an City as a new empirical research.

3.2 Data Selection and Pretreatment

Housing price has fluctuated considerably since State reform of the Chinese real estate market in 1998. Considering that housing price is always unstable due to public policy or other uncontrollable external causes, data collections were only carried out during a short time, from Mar 12, 2014 to Mar 19, 2014. All the surveyed housing samples were new-built and this will avoid the introduction of real estate depreciation factor. Moreover, using a single mouth of data (see Note 3) can avoid problems with time variations. In the end, 378-samples were taken to join the regression model. Thereby, maps of the study area including the location of the samples and administrative districts were drawn by GIS (Figure 1). With respect to spatial data, GIS software and its spatial analysis methods were taken advantage of calculating the location variables (Table 1). Basic spatial data of Xi'an City in 2013 including residential, towns, boundary, water, buildings, roads, subway stations etc. were illustrated (Figure 3a). Meanwhile housing price samples can also be visualized by GIS (Figure 3b).

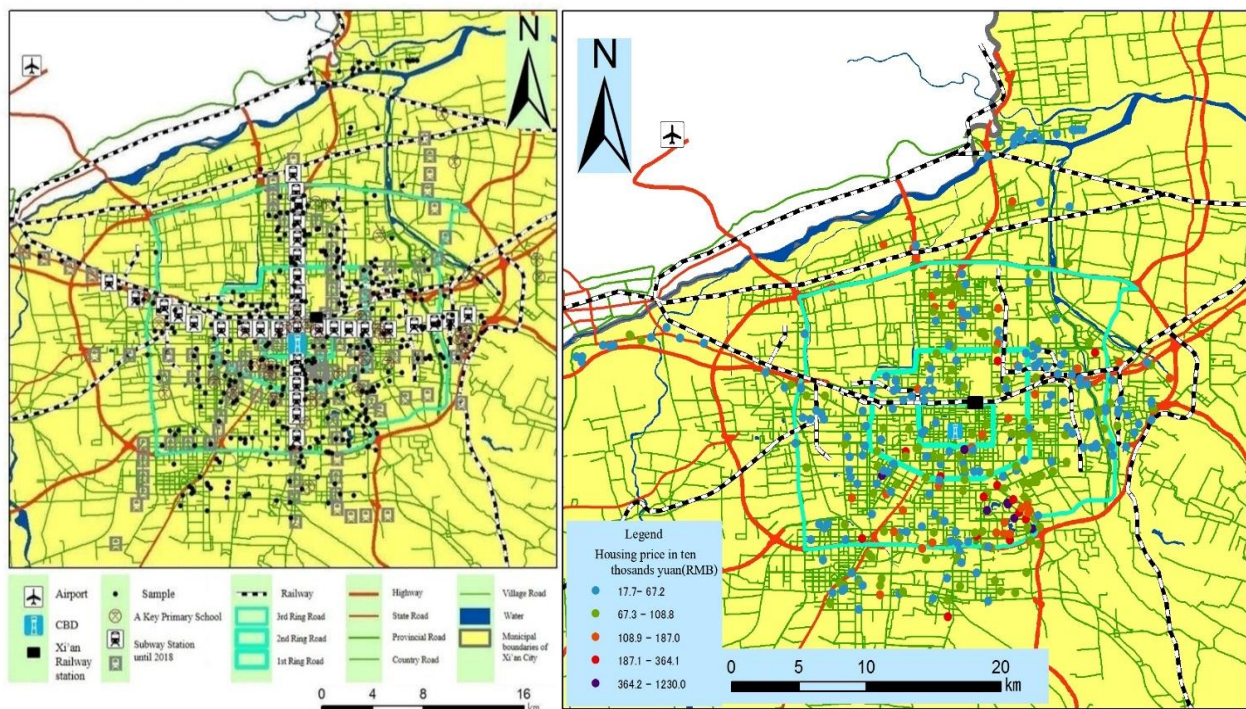


Figure 3. Spatial distribution of features in GIS (left: a) and housing price samples (right: b)

Referring to Table 1, 11 variables are shown in details. It must be noted that P denotes the dependent variable that means the total price of a new residence. Comparing with the unit-price (price per m²), it can better reflect the structural characteristics of a residence. Apparently, some hedonic literatures pointed out that property structure differs mostly with housing size and they have proved that property size has a positive relationship with selling price (McLeod, 1984). This case study attempts to find a certain relationship between property structure characteristics and total price of a new residence. In other words, the unit-price is only determined by location characteristics and neighborhood characteristics but it fails in persuasive about the relationship with the structure characteristics of a new residential property. Therefore,

I choose the total price of a new residence as the dependent variable which is added into the hedonic pricing model.

Another point is about the dummy variable, E . We need to quantify the influence of a Key Primary School in the hedonic pricing model. That is to say, we need the threshold which can retain statistical significance at least 5% level in the regression process. This threshold will be confirmed in Chapter 4.

In addition, some independent variables needed to make logarithmic transformation, which can make data distribution of every variable become smoother, that is to say, closer to follow the Normal distribution. Supported by GIS and SPSS, we had defined the following 11 variables that were transformed: Ln_DC , Ln_DSX , Ln_DS , DR , Ln_S , BR , LR , WR , HB , SN , and E .

Ln_DC denotes the logarithm of straight-line distance to the central business district. Here in particular, many scholars are partial to using road distance by network analysis instead of straight-line distance. However, calculations of road distance may introduce statistical noises and be complicated. Following previous hedonic studies (Yu, Zheng, & Liu, 2008; Wu et al., 2014; Zheng & Kahn, 2007), this study employs straight-line distance when measuring the influence effect of every characteristic. Xi'an Bell Tower is regarded as the central business district (CBD) because it has always been the city center since a long time. Ln_DSX denotes the logarithm of straight-line distance to Xi'an Railway Station. This study selected two variables that are respectively DC and DSX as explanatory variables together. According to regression results, we can decide which one can better be regarded as the final variable into the hedonic pricing model and obtain the result which is more statistically significant. Ln_DS denotes the logarithm of straight-line distance to the nearest subway station. For the sake of avoiding regression biases, this study adopt a method that every subway line was considered as the location factor when DS variable was defined, no matter whether this subway line do has been completed until April 2014 or still under construction until 2018. Roughly speaking, even these subway lines under construction can also urge the rise in prices of surrounding properties. DR denotes the distance to the nearest river or lake. Comparing with other variables, we did not take the logarithmic scale because DR variable has lower orders of magnitude. Ln_S denotes the gross floor area of the residence. BR denotes numbers of bedroom. LR denotes numbers of living-room. WR denotes numbers of washroom. HB denotes a dummy variable that means whether the residence is in a high-rise building above 12 floors or not. SN denotes a dummy variable that means whether the residence is a south facing house or not. E denotes a measure of education environment. When $E_d = 1$, where d is the radius of buffer made by a Key Primary School, there is one or more Key Primary Schools; or when $E_d = 0$, there is no Key Primary School within the d meters buffer. Here particularly, the author provisionally set $d=1000$ meters (see Note 4) and prepare to adjust this threshold which can retain statistical significance at least 5% level in the regression process.

3.3 Collinearity Diagnostics

Before the construction of hedonic pricing model, all of explanatory variables which are prepared to join the model should be further decided, because collinearity problems will make the regression coefficients unreliable. Thereby, collinearity diagnostics and stepwise regression method are applied by using SPSS. Table 2 shows the achievement of collinearity diagnostics. According to two collinearity parameters i.e. tolerance (Tol) and variance inflation factor (VIF) (see Note 5), we can take cognizance of that collinearities have been obviously existing (See italics in Table 2 left) as follows: between DC and DSX , which come from a kind of spatial correlation; among the variables: gross floor area of a residence (S) and house layout (BR , LR & WR), which come from a kind of algebraic correlation. On the other hand, to capture the chosen variables and simplify the model, stepwise regression process needs to be done repeatedly (Table 3). Specifically speaking: Step 1, remove the variables having collinearity one by one and contrast the regression effect through coefficient of determination (R^2). Step 2, retain one variable which had the highest R^2 if removing the others which exist collinearities with it. Step 3, run the stepwise regression module in SPSS software, and get better variables set which have generally eliminated collinearity problem (See Table 2 right). R^2 can be regarded as the indicator to judge whether the function form of hedonic pricing model has measured best performance with goodness-of-fit. Supported by SPSS, 7 variables had been finally chosen to join the regression model when collinearities were generally eliminated (See Table 2 right).

Table 2. Collinearity diagnostics

Variable	Start		Variable	End	
	Collinearity Tol	Statistics VIF		Collinearity Tol	Statistics VIF
<i>Ln_DC</i>	0.212	4.720	<i>Ln_DC</i>	0.638	1.567
<i>Ln_DSX</i>	0.234	4.278			
<i>Ln_DS</i>	0.697	1.434	<i>Ln_DS</i>	0.721	1.386
<i>DR</i>	0.951	1.052	<i>DR</i>	0.957	1.045
<i>Ln_S</i>	0.234	4.281	<i>Ln_S</i>	0.829	1.206
<i>BR</i>	0.311	3.215			
<i>LR</i>	0.694	1.441			
<i>WR</i>	0.422	2.367			
<i>HB</i>	0.888	1.126	<i>HB</i>	0.894	1.118
<i>SN</i>	0.824	1.213	<i>SN</i>	0.846	1.182
<i>E</i>	0.822	1.217	<i>E</i>	0.853	1.172

Table 3. Judgment process of variable selection

Variable selection in the model	Coefficient of determination (R^2)	Remarks
Every variable is added.	0.81	Start
Every variable Except <i>Ln_DC</i>	0.77	Hold <i>Ln_DC</i>
Every variable Except <i>Ln_DSX</i>	0.80	Drop <i>Ln_DSX</i>
Hold <i>Ln_S</i> and remove <i>BR,LR & WR</i>	0.78	Hold <i>Ln_S</i>
Hold <i>BR</i> and remove <i>Ln_S, LR & WR</i>	0.57	Drop <i>BR</i>
Hold <i>LR</i> and remove <i>Ln_S, BR & WR</i>	0.38	Drop <i>LR</i>
Hold <i>WR</i> and remove <i>Ln_S, BR & LR</i>	0.63	Drop <i>WR</i> and end

3.4 Focusing on a Key Primary School

As Section 3.2 announced, there is a remaining problem about variable *E*. Owing to the value of dummy variable *E* is determined by the size of d-meters buffer drawn in GIS software, the radius d needs to be done some adjustments through changing the threshold value at several times, for the purpose of quantifying the capitalization of premium from a Key Primary School in the hedonic pricing model, also must meet the condition passing in T-test at statistically significance (Table 4). Finally when setting an 825-meters buffer, we can get the best regression result with statistical significance at 5% level. This study had tried 8 times and finally found a certain threshold value, 825 meters. Results prove that, within a scope of 825-meters, Key Primary Schools, as one of the most representative local public goods, have the strongest impact on new residential property values in Xi'an City. As noted in Sections 3.3 and 3.4, this study had selected 7 variables and simplified the model successfully.

Table 4. Adjustment for the variable E

<i>E</i>	Coefficients	t-Ratio	Degree of confidence in T-text	Passed or not
<i>E_1000</i>	0.018	0.692	51.1% (<<95%, try to set d<1000)	Resolutely not passed
<i>E_800</i>	0.040	1.620	89.4% (become better, try to set d<800)	Not passed
<i>E_600</i>	0.020	-0.169	Regression failed	Resolutely not passed
<i>E_700</i>	0.027	1.079	71.9% (worse than d=800)	Not passed
<i>E_750</i>	0.025	1.024	69.4% (even worse than d=700, must set d>800)	Not passed
<i>E_900</i>	0.042	1.605	89.1% (become better, but worse than d=800, try to set 800<d<900)	Not passed
<i>E_850</i>	0.048	1.865	93.7% (become much better, try to set d=825)	Could be passed but statistically significant only at the 10% level
<i>E_825</i>	0.090	0.043	95.7%>95% (It means when d=825, variable <i>E</i> works best in hedonic pricing model)	Passed T-test at the 5% level and can be added into the regression

4. Results of Hedonic Pricing Model

Results of parameter estimates are presented in Table 5. The estimated coefficients generated using hedonic pricing model exhibits signs in accordance with expectations. They provide a clear indication that factors of housing area, CBD; floors; subway station; Key Primary School and south facing do exert a significant impact on new residential property values. Final regression equation denoting the whole results can also be created (Equation 4).

Table 5. Regression results after eliminating collinear variables, dependent variable: Ln (P), where P is total price

Independent variables	Regression parameters		
	$(R^2 = 0.782, \text{ adjusted } R^2 = 0.779, \text{ F-value} = 267^{***}, \text{ No. of Obs.} = 378)$		
	Coefficient	t-Ratio	P-value
Constant	1.413	5.139	0.000
Ln_S^{***}	1.245	32.311	0.000
Ln_DC^{***}	-0.268	-0.924	0.000
HB^{***}	-0.142	-4.538	0.000
Ln_DS^{***}	-0.061	-3.174	0.002
E^{**}	0.090	2.097	0.037
SN^*	0.044	1.690	0.092

*, ** and *** indicate statistical significance at 10%, 5% and 1% level, respectively

$$Ln(P) = 1.413 + 1.245Ln(S) - 0.268Ln(DC) - 0.142HB - 0.061Ln(DS) + 0.090E + 0.044SN \quad (4)$$

From the value of every coefficient, the calculation based on the regression equation can be completed, in order to quantify the estimation value of capitalization corresponding every variable, respectively. Premiums can be calculated according to series of processes of anti-logarithmic transformation.

Main findings are summarized as follows:

1. P (total price in 10,000 RMB) increases by 1.245% along with S (Area in m^2) increases by 1%. According to the average level of housing prices got from the 378-samples (922,500 RMB), this is equivalent to about 11,500 RMB more expensive.
2. P decreases by about 0.268% along with DC (distance to CBD in km) becomes 1% farther, which is about 2500 RMB cheaper. Generally, between a residence and CBD there exist strong distance decay effects.
3. A residence at the floors which is above 12th (≥ 12) is 14.2% cheaper than one at a lower floor below 12th (< 12). According to PCDI (per capita disposable income) of Xi'an in 2013 (see Note 6), two ratiocinations can be generalized that: 1) a high premium caused by the number of the floor; 2) home-buyers should pay more about 121,800 RMB to buy a low-rise residence, which is equivalent to nearly 3.7 times the amount of PCDI in Xi'an in 2013.
4. When distance to the nearest subway station (in km) becomes 1% farther, housing price will become 0.061% cheaper, correspondingly. It means that about 560 RMB cheaper at the average level. Comparing with other metropolis cities that have well-developed rail transportation, Xi'an is during a period of the initial stage (until 2014 at least) as a developing city because of a low-level premium (see Note 7). In contrast, Metropolis cities like Beijing, Shanghai or Guangzhou, there are much higher premium about the factor of rail transportation (Gu & Zheng, 2010). Li et al. (2011) found that inside CBD within 8km of Beijing, premiums caused by subway station is not significant and cannot be passed in regression processing, that is to say the factor of subway station failed to join the hedonic pricing model because the result of T-test is substandard. Only outside the area which is 8km far from the CBD, premiums on the housing price caused by subway station seemed to be statistically significant. This contradiction can be considered relating to the commuting cost. The metro rail transit is more inclined to give scope to its advantages about commuting cost in the suburban districts far from the CBD (Dong, Ding, & Zhao, 2009). Conversely, taking other modes of transportation like bus is cheaper and more efficient than taking subway rails for commuting activities happening around the CBD.
5. With regard to the premium from a Key Primary School, three arguments can be supported: 1) In the case that when setting an 825-meter buffer of a Key Primary School, the best regression results can be got with statistical significance at 5% level. For the future study, another approach named matching regression can be used to specifically investigate the so-called *Xue Qu Fang* problem (Hu et al., 2014); 2) a within-zone housing unit is sold 9.4% more than if it was outside the attendance zone of a Key Primary School; 3) At average, premiums are nearly 86,900 RMB from the factor, a Key Primary School, which is about 2.6 times the amount of 2013 year's PCDI of Xi'an. It also indicates that educational resources have been well capitalized into the residential property value.
6. For the factor about the situation of North and South transparent, there is a slight premium, which is only 41,500 RMB. Moreover, the regression result is roughly statistically significance at only 10% level (strictly speaking it is normal at least 5% in statistics). This is because the vast majority of new residences are south facing, which can be clearly read from the raw data. Therefore, it seems that data set are lack of comparison samples (not N-S facing houses) when just discussing about new residential property values.
7. After T-text, all of these chosen variables in the hedonic pricing model yield good explanatory power and can also reflect the preferences of homeowners in Xi'an City. However, during the regression process, there is a meaningless variable, DR , which was dropped in the regression process. In general, housing price tends to become higher when the residence is closer to a river or lake according to the expectation. But the regression coefficient of DR got a positive number, which is out of expected results and the actual situation. Perhaps, this situation is related to the limitation of the

samples, or else unobserved variables or errors. Thence, this study has proved a mathematical corollary that omitted variables may result in regression failure and estimated bias. Besides, water system in Xi'an City is very complicated, which may cause geographical recognition errors, such as the Xi'an Moat surrounding the City Wall and the *Weihe* Reservoir. So nearest neighborhood analysis done by GIS may also cause accidental and systematic errors. Additionally, for these new residential properties which located in the undeveloped suburban areas in Xi'an, the impacts of water on housing price are relatively weak, even though they are near a river or lake with a very low price but a closed distance to some suburban rivers (see Note 8). Furthermore, real estate development is faster but development of urban infrastructure lags behind. So this factor failed to join and estimate the regression equation. All of these above situations could not be controlled; consequently, it is necessary to employ the approach of residual analysis, which will be brought to light in Chapter 5.

5. Residual Analysis

According to the above work, firstly the reasons why residual analysis is necessary need to be discussed:

1. Variable *DR* failed to join the hedonic pricing model and was dropped by using stepwise regression method, caused by spatial local singularities or data limitations.
2. Unobserved variables are existent. Quantity of residual can reflect the effectiveness of regression fitting, as well as residuals are related to the existence of unobserved variables in a sense.
3. Positive or negative numbers of residuals indicate smaller or larger biased estimate of regression equation, respectively.
4. It provides a reference for future studies using residual analysis. Specifically, through the process of residual analysis, we can try to explain what places have more caused residuals and the reason in consideration of city layout of Xi'an.
5. According to the results running in SPSS software, we can probe the local singularity from a table of residuals. Distribution of residuals was approximately following the Normal distribution, and this indicated that the chosen variables in the hedonic pricing model yielded good explanatory power and residuals were under a reasonable situation. Herein, map of the location of residuals can be illustrated in Figure 4a. Moreover, sample cases with a high residual are picked out and visualized by deviation multiples (Predictive values / observed values), can used for reflecting and quantifying the local singularities (Figure 4b).

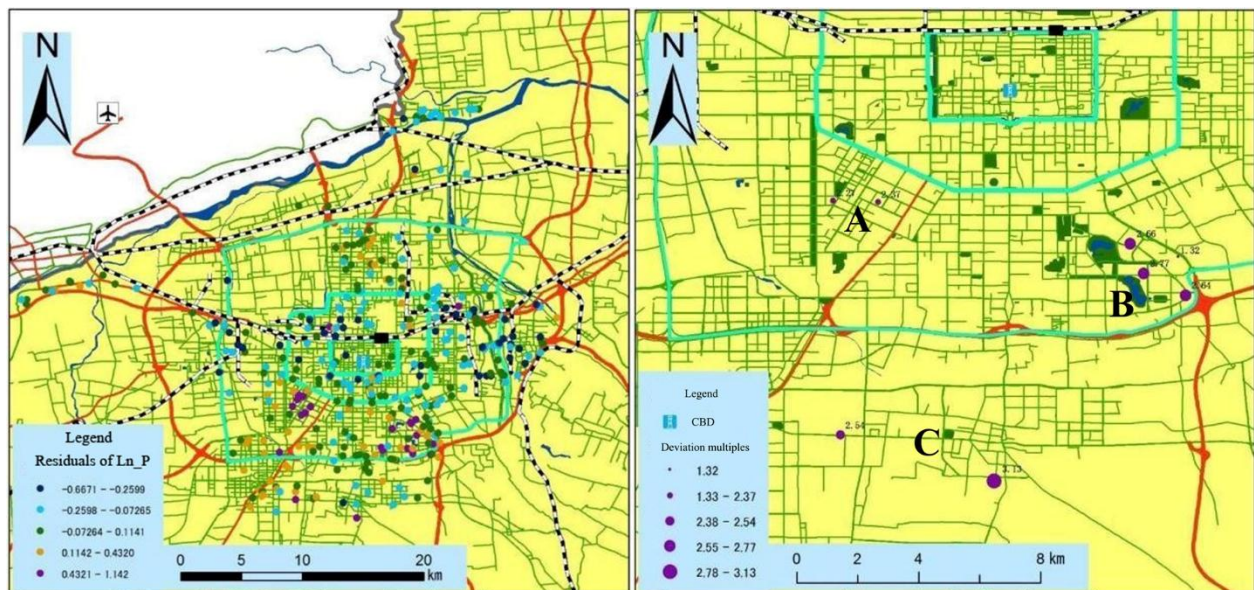


Figure 4. Spatial distribution of regression residuals (left: a) and local singularities (right: b)

On Figure 4a, we can find that most of residuals are located in the developing area between 2nd Ring Road and 3rd Ring Road, which have more real estate development and appreciation space. Even though these areas are about 4km-8km far from the CBD, more and more new residences have been built in recent years. In other words, in these places effects of unobserved variables on the housing price are more significant. So, the results of hedonic pricing model show a situation of smaller biased estimates. Eight of the local singularity points that have the highest residual is illustrated in Fig. 4b. They are located in three areas in Xi'an City. First area (A) in the southwest of the 2nd Ring Road, High-tech New Development Zone is located. It seems that homeowners in High-tech New Development Zone can

enjoy better educational and commercial resources. Second area (B), *Qujiang* District is known as the *Qujiang* scenic spot, which is characterized by its style of Tang Dynasty and has been the most famous royal park and natural scenic area in Chinese history. In the respect of samples, many new residences in *Qujiang* District were built as villas and International communities. Third place (C) is far-away from the 3rd Ring Road in *Chang'an* District, which is known as a College Town. Here we can make a forecast on real estate development in Xi'an: in the future, housing price in *Chang'an* District may increase, which is reflecting the developed prospects as a new downtown. All of the aforementioned residuals are smaller biased estimators, which is because premium caused by other factors exceeded the attenuation effects of distance to CBD. It is obvious that unobserved variables induce more bias in the regions with high residuals (including positive and negative). Furthermore, another corollary could be drawn that larger biased estimators are located in *Bahe* District and western Xi'an City (Fig. 4a). The reasons can be explained that: some subway lines have not been yet operational in *Bahe* until 2018, probably; large numbers of low-cost housings are located in these areas. Real estate development of these places is backward comparing with other places in Xi'an City.

To many real estate investors and planners, it is important to grasp the trend of housing market. Herein, this study put forward a hypothesis that price-rise of housing in these areas would be evident continuing to the future. In order to inquiry and prove this hypothesis, this study tried to apply to another approach of cluster analysis in Chapter 6.

6. Regional Classification by Using Cluster Analysis

Apart from focusing on micro-Geographical matters, it is necessary to grasp the overall situation of new real estate market in Xi'an City by using sample data. Another approach of cluster analysis is employed in this paper, which is widely used in geography to make spatial classification. In other words, the regional development trend and differences could be exposed by reading the results of classification. Herein, this study made a survey on the whole real estate market of Xi'an City in a particular month, considering several indices in 12 regions divided by real estate market. K-means method was employed supported by SPSS. Figure 5 shows the dendrogram constructed by SPSS, which can be used to determine the number of clusters and the characteristics of each cluster.

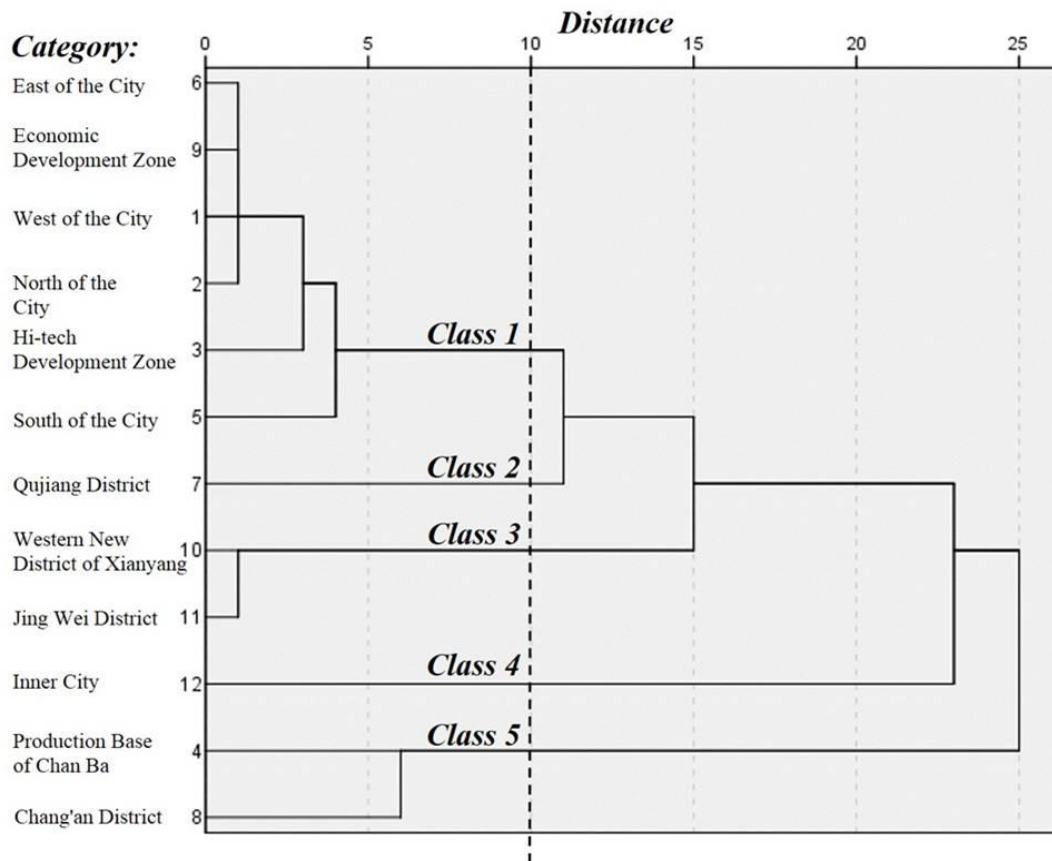


Figure 5. Result of cluster analysis

Details in the Figure 5 can be described as follows:

1. Class 1: East of the City, Economic Development Zone, Hi-tech Development Zone, West of the City, North of the

City, South of the City.

6 of 12 districts were divided into the first class, which have the following characteristics: housing price at a high-medium level; low growth rates even negative growth; quite larger amount of new residences; located within the 2nd Ring Road.

2. Class 2: *Qujiang* District

Qujiang district was divided into the second class, specially. The characteristics can be described as: housing price at an extremely high level; substantial growth rate; quite larger amount of new residences.

3. Class 3: Western New District of *Xian Yang*, *Jing Wei* District

Two districts far away located in outer suburbs, were divided into the third class. The characteristics are shown in details: housing price at a low level; no-growth rate; small amount of new residences;

4. Class 4: Inner City

One district next to the CBD can be divided into Class 4, where has a quite high housing price and minority amount.

5. Class 5. Production Base of *Chan Ba*, *Chang'an* District

Two districts located in the south suburbs of Xi'an City were divided into Class 5. Residences in these two districts have a medium-level housing price but with a highest growth rate.

In summary, real estate industry is developing in several specific districts. Overall, hypotheses put forward in the last Section can be testified—in south and east suburbs, the development of real estate market is more remarkable. As result shows, five housing agglomerations have been classified in Xi'an, depending on their average price per m², numbers of new residences and monthly growth rate in different regions. Furthermore, we can reveal the varying trends and prospects on the real estate market, by reading local disproportions and characteristics of housing price through the cluster analysis.

7. Discussions and Conclusions

This study reported an effective attempt to apply hedonic pricing models to reveal influencing factors and spatial characteristics of the housing price, and highlights the fact that hedonic pricing models would be improved by embodying spatial characteristic variables captured by GIS and data processing supported by SPSS; it thus contributes to the further joining of geography, statistics and economics, and promotes development of urban economics. Results indicate that:

1. A within-zone housing unit is sold 9.4% more than if it was outside the attendance zone of a Key Primary School. Besides, this study educed a specific threshold value (825 meters) about the residence inside a Key Primary School's attendance zone in several repeating regression processes in order to achieve the target that statistical significance at the 5% level. Herein, the premium on residential properties from the influence of educational resources has been quantized, i.e. home buyers have how much willingness-to-pay for the high-quality educational resource.
2. A corollary was got that advantages of rail transportation had not yet adequately been brought into full play in Xi'an City, and this detail is different like other hedonic approach.
3. Unlike previous studies which just unilaterally used hedonic price modeling but ignored residuals problem, residual analysis and cluster analysis are applied to understand the local differences on housing pricing and the development situation of real estate market in Xi'an City. This can be seen as a kind of innovation, which combines the regional differences (coming from real world) and the spatial distributions of singularities (feedback from data). Through these further multivariable analyses after finishing the hedonic regression, this study attempts to further reveal the spatial singularity and local characteristics. Three special areas could be found regarded as high-technology area (A), luxury residential area (B), and new town (C) in Xi'an City. Moreover, results of cluster analysis indicated that real estate of Xi'an City was developing vigorously in the east and south of suburbs. In conclusion, this paper successfully assessed the housing price trends and the real estate market prospects for new residences in Xi'an in a new way different from recent literatures.
4. As the case of this study about hedonic pricing modeling, the utilization of GIS and SPSS is demonstrated to be effective and efficient.

On the other hand, some limitations are inevitably existent. This study investigates only the total value of a new residence, whereas secondary housing and time series problems were both avoided. These omission variables might make the process complicated and suffer from regression bias. Thereby, it is possible that some variables (*DR*) fails to join the hedonic pricing model. However, results of 6 variables are consistent with expectations, and this suggests that collinearity diagnostics and stepwise regression method is effective in hedonic approach.

Another helpful precious study using another approach named matching regression can be regarded as a good reference, which specifically investigated the *Xue Qu Fang* problem (Hu et al., 2014). But this study just inherits the main idea of this paper and formed mutual corroboration in a certain sense, whereas did not consider on this issue in depth. Apparently, effects of a Key Primary School on housing price are positive, and premium has also been correctly estimated through hedonic pricing model.

In consequence, it is hoped that the results of this study will provide a reference for research on housing price applied hedonic pricing model.

Note

1. In Chinese cities only property owners rather than renters are entitled to enroll in local public school when both live in a school's attendance zone.
2. *Xue Qu Fang* is a Chinese word. Its English phrase is school district housing. In China it specifically refers to the residences which are located inside a Key Primary School's attendance zone. It is also recognized as a hot vocabulary by the public but hard to be strictly defined by academic world.
3. In Chinese real estate market, housing price survey and report from official is usually updated once a month.
4. In order to ascertain an exact value which can best fit hedonic regression, an initial value must be required.
5. In statistics, the bigger *VIF* is, or the smaller *Tol* is, the more collinearities there are.
6. Annual per capita disposable income of Xi'an City in 2013 was 33,100 RMB (Shaanxi Provincial Bureau of Statistics, 2013).
7. The first subway line in Xi'an was opening to traffic in 2011. Considering the current data in 2013, there are only two subway lines No.1 and No.2 crossing the central East-west and South-north axes of Xi'an City. However, several other subway lines are under construction and will be excepting to be operational until 2018. Hereby, this study selected 4 subway lines for their long-term and potential capitalization effects.
8. There are three big rivers, the *Weihe* River, the *Bahe* River and *Chanhe* River located in the suburban areas in Xi'an City. They are far from CBD, and many housing price samples are located nearby.

Acknowledgement

This work was carried out by the illuminating instruction of my supervisor, Prof. Yuji Murayama. I also owe my sincere gratitude to my fellowclassmates.

References

- Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22, 453-474. [https://doi.org/10.1016/0166-0462\(92\)90039-4](https://doi.org/10.1016/0166-0462(92)90039-4)
- Cassel, E., & Mendelsohn, R. (1985). The choice of functional forms for hedonic price equations: comment. *Journal of Urban Economics*, 18, 135-142. [https://doi.org/10.1016/0094-1190\(85\)90012-9](https://doi.org/10.1016/0094-1190(85)90012-9)
- Cook, R. D., & Weisberg, S. (1982). *Residuals and Influence in Regression*. New York: Chapman and Hall.
- Ding, W. J., Zheng, S. Q., & Guo, X. Y. (2010). Value of access to jobs and amenities: Evidence from new residential properties in Beijing. *Tsinghua Science & Technology*, 15, 595-603. [https://doi.org/10.1016/s1007-0214\(10\)70106-1](https://doi.org/10.1016/s1007-0214(10)70106-1)
- Dong, F., Ding, H., & Zhao, A. P. (2009). Study on the affecting factors of housing prices along rail transit based on hedonic price model - case study of Beijing Subway Line 5 [in Chinese]. *Journal of Beijing Normal University (Social Science)*, 4, 137-143. <https://doi.org/10.3969/j.issn.1002-0209.2009.04.016>
- Freeman III, A.M. (1979). *Benefits of environmental improvements: theory and practice*. Baltimore MD: Johns Hopkins University Press/Resources for the Future.
- Gu, Y. Z., & Zheng, S. Q. (2010). The impacts of rail transit on property values and land development intensity: The case of No. 13 line in Beijing. *Acta Geographica Sinica*, 2, 213-223. <https://doi.org/10.2139/ssrn.1150058>
- Haas, G. C. (1922a). A statistical analysis of farm sales in blue earth county, Minnesota, as a basis for farm land appraisal. (Unpublished Master's thesis). The University of Minnesota, Minnesota, USA
- Haas, G. C. (1922b). Sale prices as a basis for farm land appraisal. *Technical Bulletin 9, St. Paul*: The University of Minnesota Agricultural Experiment Station.
- Halvorsen, R., & Pollakowski, H. O. (1981). Choice of function form for hedonic price equations. *Journal of Urban Economics*, 10, 37-49. [https://doi.org/10.1016/0094-1190\(81\)90021-8](https://doi.org/10.1016/0094-1190(81)90021-8)

- Harrison, D., & Rubinfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management*, 5, 81-102. [https://doi.org/10.1016/0095-0696\(78\)90006-2](https://doi.org/10.1016/0095-0696(78)90006-2)
- Hu, W. Y., Zheng, S. Q., & Wang, R. (2014). The capitalization of school quality in home value: A matching regression approach with housing price-rent comparison [in Chinese]. *China Economic Quarterly*, 13, 384-403. <https://doi.org/10.13821/j.cnki.ceq.2014.03.017>
- Judd, G. D., & Watts, J. M. (1981). Schools and Housing Values. *Land Economics*, 57, 459-470. <https://doi.org/10.2307/3146025>
- Kim, J., & Zhang, M. (2005). Determining transit's impact on Seoul commercial land values. *International Real Estate Review*, 8, 1-26. Retrieved from <http://EconPapers.repec.org/RePEc:ire:issued:v:08:n:01:2005:p:1-26>
- Kong, F., Yin, H., & Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning*, 79, 240-252. <https://doi.org/10.1016/j.landurbplan.2006.02.013>
- Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74, 132-56. https://doi.org/10.1007/978-3-642-51565-1_34
- Li, G., Huang, X. Y., Liao, S. L., Han, F. Q., & Yuan, R. D. (2005) Research on algorithms of grey residual error model in highway traffic accidents forecast. *Wavelet Analysis and Active Media Technology* [Internet]. World Scientific Publishing Company; 675-682. Retrieved from http://dx.doi.org/10.1142/9789812701695_0101
- Li, X. P., Silva, E. A., & Bo, S. X. (2011). A case study on the hedonic price model in Beijing City [In Chinese]. *China Market*, 42, 58-63. <https://doi.org/10.3969/j.issn.1005-6432.2011.42.012>
- Liu, R. Q., & Sun, X. Y. (2015). The effect of education quality capitalization on housing prices based on empirical analysis of elementary school district of Wuhou district in Chengdu. *Finance & Economics*, 8, 009. <https://doi.org/10.3969/j.issn.1000-8306.2015.08.009>
- Maurer, R., Pitzer, M., & Sebastian, S. (2004). Hedonic price indices for the Paris housing market. *Allgemeines Statistisches Archiv*, 88, 303-326. <https://doi.org/10.1007/s101820400173>
- McLeod, P. B. (1984). The demand for local amenity: a hedonic price analysis. *Environment and Planning A*, 16, 389-400. <https://doi.org/10.1068/a160389>
- Morancho, A. B. (2003). A hedonic valuation of urban green areas. *Landscape and Urban Planning*, 66, 35-41. [https://doi.org/10.1016/s0169-2046\(03\)00093-8](https://doi.org/10.1016/s0169-2046(03)00093-8)
- Muth, R. F. (1966). Household production and consumer demand functions. *Econometrica: Journal of the Econometric Society*, 34, 699-708. <https://doi.org/10.2307/1909778>
- Nguy, A., Sun, C., & Zheng, S. Q. (2014). Airport noise and residential property values: evidence from Beijing. In *Proceedings of the 17th International Symposium on Advancement of Construction Management and Real Estate*, (pp. 473-481). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35548-6_49
- O'sullivan, A. (2007). *Urban economics*. McGraw-Hill/Irwin.
- Pan, H. X., Ren, C. Y., & Yang, T. Y. (2007). A study on the impact on the land use of station areas brought by urban rail transport in Shanghai [in Chinese]. *Urban Planning Forum*, 4, 92-97. <https://doi.org/10.3969/j.issn.1000-3363.2007.04.013>
- Ridker, R. G., & Henning, J. A. (1967). The determinants of residential property values with special reference to air pollution. *The Review of Economics and Statistics*, 49, 246-257. <https://doi.org/10.2307/19282>
- Rosen, H. S., & Fullerton, D. J. (1977). A Note on Local Tax Rates, Public Benefit Levels and Property Values. *Journal of Political Economy*, 85, 433-440. <https://doi.org/10.1086/260575>
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82, 34-55. <https://doi.org/10.1086/260169>
- Shaanxi Provincial Bureau of Statistics. (2013). *Shaanxi Provincial Yearbook* [In Chinese]. China Statistics Press, Beijing.
- Shi, Y. S., & Wang, Y. T. (2014) The Impacting Mechanism of Housing Prices in the School Districts in Shanghai City [In Chinese]. *China Land Science*, 28, 47-55. <https://doi.org/10.3969/j.issn.1001-8158.2014.12.007>
- Shultz, S. D., & King, D. A. (2001). The use of census data for hedonic price estimates of open-space amenities and land use. *The Journal of Real Estate Finance and Economics*, 22, 239-252. Retrieved from

<https://link.springer.com/article/10.1023/A:1007895631071>

- Tiebout, C. M. (1956). A Pure Theory of Local Expenditure. *Journal of Political Economy*, 64, 416-24. Retrieved from <http://www.jstor.org/stable/1826343>
- Wang, X., Ge, Y. S., & Zhang H. (2010). School district housing's price mechanism in Nanjing Old Town [In Chinese]. *CO-Operative Economy & Science*, 12, 10-13. <https://doi.org/10.3969/j.issn.1672-190X.2010.12.004>
- Wang, X., Zhu, D. L., & Zhang, M. M. (2004). Analysis of the influence of rail transit on the distribution layout of real estate prices — a case study of Beijing light rail line No.13 [in Chinese]. *Urban Problems*, 6, 39-42. <https://doi.org/10.3969/j.issn.1002-2031.2004.06.010>
- Wei, J. F., Wei, Y. G., & Jiang, K. (2014) Study on the Affecting Factors of Housing Prices along Rail Transit Based on Hedonic Price Model-Case Study of Beijing Subway Line 5. *Applied Mechanics and Materials*, 507,642-645. <https://doi.org/10.4028/www.scientific.net/amm.507.642>
- Wen, H. Z., & Jia, S. H. (2004). A new method of real estate evaluation based on hedonic pricing model [in Chinese]. *Foreign Economies & Management*, 6, 31-35, 40. <https://doi.org/10.3969/j.issn.1001-4950.2004.06.006>
- Wu, J., Deng, Y. H., & Liu, H. Y. (2014). House price index construction in the nascent housing market: the case of China. *The Journal of Real Estate Finance and Economics*, 48, 522-545. <https://doi.org/10.1007/s11146-013-9416-1>
- Xu, A. P., Gui, Q. M., & Han, S. H. (2008). The comparison of three residuals in Kalman filtering model [in Chinese]. *Journal of Geodesy and Geodynamics*, 28,101-104. <https://doi.org/10.3969/j.issn.1671-5942.2008.01.021>
- Yu, L., Zheng, S. Q., & Liu, H. Y. (2008). The spatial variation and affecting factors of the housing price gradients – the case study of Beijing (In Chinese). *Jingji dili [Economic Geography]* 28, 406-410. Retrieved from http://en.cnki.com.cn/Article_en/CJFDTotal-JJDL200803015.htm
- Zhang, Y. J., Zheng, S. Q., & Sun, C. (2014). What's the effect of urban villages on commercial housing price? An analysis based on second-hand housing transactions in Beijing. In *Proceedings of the 17th International Symposium on Advancement of Construction Management and Real Estate*, (pp. 173-183). Springer Berlin Heidelberg.
- Zheng, S. Q., & Kahn, M. E. (2007). Land and residential property markets in a booming economy: new evidence from Beijing. *Journal of Urban Economics*, 63, 743-757. <https://doi.org/10.1016/j.jue.2007.04.010>
- Zheng, S. Q., & Kahn, M. E. (2013). Does government investment in local public goods spur gentrification? Evidence from Beijing. *Real Estate Economics*, 41, 1-28. <https://doi.org/10.1111/j.1540-6229.2012.00339>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the [Creative Commons Attribution license](#) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.