Journal of Economics and Public Finance
ISSN 2377-1038 (Print) ISSN 2377-1046 (Online)
Vol. 9, No. 2, 2023
www.scholink.org/ojs/index.php/jepf

Original Paper

The Formation Mechanism of House Price Differences in China's Urban Area in the Same City —Based on Lasso Bayesian Model Averaging Method

Jiayao Pan[1,2] & Shaoling Ding[1,2*]

[1] College of Science, Guilin University of Technology, Guilin 541004, China

[2] Guangxi Colleges and Universities Key Laboratory of Applied Statistics, Guilin, 541004, China

[*] Corresponding author.

Received: March 13, 2022 Accepted: March 22, 2022 Online Published: March 23, 2022

doi:10.22158/jepf.v9n2p48 URL: http://dx.doi.org/10.22158/jepf.v9n2p48

Abstract

With the development of economy and the expansion of urban scale, the variation of housing prices within Chinese cities has gradually become significant. Based on the housing prices of 79 districts in Chengdu, this paper analyzes the reasons for such differences. With an index system of the price difference of locational housing, a semi-log linear regression Lasso Bayesian average model is constructed, which can realize variable selection while estimating coefficients. Empirical evidence shows that the inequality of social-economic resources and the imbalance of housing supply and demand in the area are important internal factors that lead to the difference in housing prices between areas within the city; the related housing prices have an extremely significant positive impact, showing that the mobility of the place of purchase and the contagion of the house price are the main reasons for housing price; in addition, influential factors have strong positive interaction effects.

Keywords

Housing price difference, Indicator system, Hedonic Model, Lasso, Bayesian Model Averaging Method

1. Introduction

The rapid growth of housing prices in China's cities in recent years has led to a more pronounced market differentiation among different areas within cities (Bangura & Lee, 2019). To a certain extent, the market differentiation of house prices in inner cities reflects differences in the level of public services, resource allocation and economic development among inner-city districts, and also accurately reflects the residents' willingness to pay for the value of urban districts. Therefore, the problem of localized overheating of intra-city residential prices not only attracts the attention of experts and scholars, but also arouses heated discussions among ordinary people.

Existing research on housing prices in China mostly takes provincial or 35 large and medium-sized cities as the unit of consideration (Guan & Gao, 2018; He, 2018; Wu, Deng, & Liu, 2014). Because of their overly large ranges of indicators, their representativeness is poor, which affects the interpretation and cause the lack practicality for real estate regulation policies. Although studies have investigated the diffusion of housing prices among regional cities, few have considered the chain reaction of housing submarkets within large cities (Bangura & Lee, 2019). A unified real estate market can be considered as a series of subdivisions with linkages between the segments (Goodman & Thibodeau, 1998). The segmentation of the housing market is because of spatial differences in structural characteristics, neighboring amenities, or some combination of both. In different areas of the same city, the real estate markets have certain differences due to variations in their features, making them heterogeneous and difficult to conform to the actual situation of each market by a single model. Therefore, this paper establishes corresponding hedonic models for each market segment separately and introduces the interactions among variables to represent market interactions. According to the internal relationship between each variable, we determine whether there is an interaction effect between them.

The hedonic model is used in the real estate, making up for the deficiencies of traditional real estate evaluation methods that are difficult to be measured by econometric models (Niu & Liu, 2017; Lindahl-Kiessling & Landberg, 1995; Yang, Fan, & Zhao, 2018; Zhu & Zhang, 2020; Hu, Xiong, Cai, & Yuan, 2020). It also takes consumer housing demand into the model. The theory argues that housing has a variety of different characteristic attributes, and its price is reflected by the implied price of each attribute, which is suitable for addressing the prices of heterogeneous commodities such as real estate. After a theoretical analysis of the internal relationship among location attributes, market relations and housing prices, this paper constructs an index system for housing prices in a region, and model them as feature attributes of the hedonic method. However, the housing price will be different because of the values of each feature attribute and the attribute combinations of the overall model. The multicollinearity among each feature attribute also affects the comprehensibility of the model. To overcome the defects of the hedonic model in feature selection, this paper uses Lasso to screen the variables by adjusting the penalty. Then use Bayes Model Averaging to incorporate the indeterministic of the model and variables into the model consideration, and analyze the influence of each feature variable on housing prices according to coefficients and posterior probability.

In order to alleviate the problem of unbalanced and insufficient local development of the urban housing market. Alleviate the contradiction between the people's growing needs for a better life and the unbalanced market development in the urban housing market.

Taking Chengdu as an example, this paper conducts an empirical analysis and put forward relevant suggestions to ease the contradiction between housing supply-demand and differentiation pattern of the real estate market in the city. These suggestions can also address the problem of unbalanced and insufficient local development of the urban housing market, and ease the contradiction between the people's growing needs for a better life and the unbalanced housing market development.

2. Theoretical Analysis and Index Design

The attributes of the inner-city include three aspects: humanities, society and economy, which complement each other. The humanistic and social development in the urban area affects the convenience and livability of houses. Therefore, the matching degree of humanistic-social of the area and the price of housing is the internal factor for housing demanders to buy houses. Public facility is one of the important manifestations of humanistic-social values in the area. It reflects the various needs of residents, and its allocation affects urban housing prices (Stirk-Wang & Thorsnes, 2017). The supply of public services, including the number of subways, education development, commercialization, and the quality of the environment, reflects the value of the humanistic-social attributes of the area in housing prices to a certain extent, whose uneven distribution will exacerbate differentiation.

Table 1. Location Value Index System and Description

Indicator layer		Variable	Variable data	Variable symbol	Variable description
Location Value	Human and social attributes	Number of subways	Number of subway lines in this area	ra	Reflects accessibility
		Education Level	Number of Brand elementary schools in the area	sc	Reflects education level
		Commercialization level	The large centralized commercial area in this area($>2km^2$)	bu	Reflects convenience of life and commercialization
		Park value	Large park area in this area	pa	Reflects environmental quality and livability
		Water value	The medium-large water area in this area	wa	Reflects environmental quality and livability
	Economic attributes	Resident population	Residential area for sale in this area for more than 5 years	pd	Reflecting the economic prosperity of the area and the number of potential buyers of second homes
		Industrial population	Office space in this area	oia	Reflects the economic scale of the area and the number of people with purchasing power

Urban areas have the same macroeconomic background. Resident population and industrial population of the area reflect the economic conditions. Resident population reflects the economic prosperity and also reflects the size of the potential demanders of housing (Zhu, Füss, & Rottke, 2013). Industrial population reflects the economic scale and also reflects the size of the population with potential housing purchasing ability. Both of them reflect the rigid demand-side value of housing.

According to the results of the above theoretical analysis, the design and measurement methods of location value indicators are shown in Table 1.

Market relationship reflects the activity of the residential market, the commodity nature of housing and the duality of consumption and investment in housing (Sah, Conroy, & Narwold, 2016). The market relationship of the residential market in the area is divided into two aspects: one is the relationship between supply and demand, and the other is the market environment. The relationship between supply and demand is reflected by transaction volume, inventory, and inventory-sales ratio; the market environment is reflected by brand projects, new brand projects, and land prices. Among them, brand projects and new brand projects reflect the scale of housing supply; land prices determine housing costs and housing supply. The evidence shows that the market environment reflects speculative demand in the housing market to a certain extent.

According to the results of the above theoretical analysis, the design and measurement methods of supply-demand indicators are shown in Table 2.

Table 2. Supply-demand Index System and Description

Indicator layer		Variable	Variable data	Variable symbol	Variable description
Supply-demand	Supply - demand	Sales	The area of the board in this area that was sold in the month	*sa*	Reflects residential market volume and supply-demand
		Stock	The area of the board in this area that was not sold in the month	*st*	
		Stock-to-sales ratio	Sum of stock sales in this area divided by sales	*ssr*	
	Market Environment	Brand projects	Cumulative area of brand corporate projects in this area	*br*	Reflects the scale of residential supply
		New brand	New brand corporate project area for the month	*nbr*	
		Land price	The price of land in this area in the current month or previous land transactions	*lp*	Affects the cost and the availability of housing

With the increase of capital, labor and information flows, the urban real estate market is increasingly linked (Panduro & Thorsen, 2014). Based on geographical factors and real estate industry experience, geographically adjacent areas or areas with close real estate market connections are defined as related areas (Tibshirani, 1996; Xin & Khalid, 2018; Fan & Li, 2001). Housing price changes in one area depend on that in other areas (Panduro & Thorsen, 2014), so the housing market in related areas has an important impact on residential prices in this area. In addition, the influencing factors of housing prices are interrelated and complex, and there are superimposed spillover effects between the humanistic, social, economic attributes and market relations. Thus the interaction effects between factors must also be considered.

According to the results of the above theoretical analysis, the design and measurement methods of associated area impact indicators are shown in Table 3.

Table 3. Associated Area Impact Index System and Description

Indicator layer	Variable	Variable data	Variable symbol	Variable description
Associated Area Impact	Associated resident population	Residential area in the associated area sold for more than 5 years	*lpd*	The impact of the market demand side of the associated residential market on residential prices in this area
	Associated industrial population	Office space in associated areas	*loia*	
	Affiliate stock-to-sales ratio	The sum of the inventory sales of the associated area divided by the sales volume	*lssr*	Reflecting the proximity spillover effect of the residential market and spatial radiosity
	Associated residential price	The price of land transactions in the current month or previous land transactions	*lhp*	

3. Semi-log Linear Regression Lasso Bayesian Model Averaging

This paper constructs a new semi-log-linear regression Lasso Bayesian model averaging. The model is based on the semi-logarithmic model in Hedonic theory, and decides whether to use the logarithmic form according to the value of the independent variable. Then uses Lasso to select the feature variable. Finally, this paper uses Bayes Model Averaging to consider the non-deterministic nature of the model itself to get the final model result. Different from the general multiple regression model, this method can automatically compress the regression coefficient of the independent variable with less influence to 0 to a greater extent. Thus realize the selection of variables while estimating the coefficient. The steps are:

First, use Hedonic (Sah, Conroy & Narwold, 2016) for modeling. Residential prices are formed by various factors and features, so the combinations of each feature are various. The hedonic model not only has a complete theoretical foundation but also has simple and easy-to-understand basic functions. This paper changes its form to explain the form of different residential prices more effectively. The functional forms of the hedonic model are expressed as three types: linear equations, logarithmic equations and semi-logarithmic equations:

linear：

$$P_i = \alpha_0 + \sum_{t=1}^{n} \alpha_t x_{ti} + \varepsilon_i \tag{1}$$

log-linear：

$$\ln P_i = \alpha_0 + \sum_{t=1}^{n} \alpha_t x_{ti} + \varepsilon_i \tag{2}$$

log-log：

$$\ln P_i = \alpha_0 + \sum_{t=1}^{n} \alpha_t \ln x_{ti} + \varepsilon_i \tag{3}$$

Among them: P_i is the price of the i^{th} residential area; x_{ti} is the t^{th} attribute of the i^{th} residential area; ε_i is the stochastic error term of the model; α_0 is the constant term to be estimated.

In the using process, the parameters of the linear model are often not realistic, and the full logarithmic model cannot intuitively explain the actual market situation. In contrast, the semi-logarithmic form is more realistic. However, because of the different data types and the large magnitude gap among the variables, the simple semi-logarithmic form model cannot achieve the expected effect. Therefore, the variables are processed according to the size of their value, as shown in Table 4.

Table 4. Variable Format

Processing method	Variable symbol
Logarithm	pd , oia , sa , st , lp , lpd , $loia$, lhp
Linear	ra , sc , bu , ssr , $lssr$
According to the size of the value, determine whether to take the logarithm	pa , wa , br , nbr

After processing each variable, a semi-logarithmic model based on the hedonic method is obtained:

$$\ln P_i = \alpha_i + \sum_{j=1}^{n_1} \beta_j \ln x_{ij} + \sum_{m=1}^{n_2} r_m \ln x_{im} + \varepsilon_i \tag{4}$$

Where P_i is the average price of housing in the i^{th} area; $\ln x_{ij}$ is each variable that takes the logarithm; x_{im} is the variable that maintains the linear prototype; α_i is a constant term; $\beta_1,\cdots,\beta_{n_1}$, $\gamma_1,\cdots,\gamma_{n_2}$ are regression coefficients; ε_i is a stochastic interference term; n_1 is the number of independent variables in the form of logarithms in the area; n_2 is the number of independent variables that maintain the linear prototype; since the total data set is 79 areas in Chengdu Residential prices, so $i=1,2,\cdots,79$.

And because the homogeneous nature of the real estate market is one of the basic assumptions of the hedonic theory, the analyzed market required to be a unified market. However, in different areas of the same city, the real estate market has certain differences because of changes in its influencing factors. Thus the basic assumption of market homogeneous is fail. It can be considered that the real estate market in the same city is composed of a series of market segments, with different hedonic functions. That is, according to the different variable forms of park value, water value, brand projects and new brand projects, the number of and in the semi-logarithmic model of housing price in each area is not the same.

Because there are interaction effects among the influencing factors, selecting one or several of the variables cannot accurately predict housing price. Based on the internal relationship, this paper includes the interaction effect into the model. Modify formula (4) to get formula (5).

$$\begin{aligned}\ln P_i &= \alpha_i + \sum_{j=1}^{n_1}\beta_j \ln x_{ij} + \sum_{m=1}^{n_2} r_m \ln x_{im} + \sum_{k=1}^{n_3}\lambda_k x_{im} \ln x_{ij} \\ &+ \sum_{p=1}^{n_4}\eta_p x_{im_1} x_{im_2} + \sum_{q=1}^{n_5}\theta_q \ln x_{ij_1} \ln x_{ij_2} + \varepsilon_i\end{aligned} \tag{5}$$

$\lambda_1,\cdots,\lambda_{n_3},\eta_1,\cdots,\eta_{n_4},\theta_q,\cdots,\theta_{n_5}$ are regression coefficients, their corresponding terms represent the interaction of each variable, whose variable forms are the same as those in Table 4; $j=1,2,\cdots,n_1$; $m=1,2,\cdots,n_2$; $k=1,2,\cdots,n_3$; $p=1,2,\cdots,n_4$; $q=1,2,\cdots,n_5$; $m_1 \neq m_2$; $j_1 \neq j_2$.

Housing price is a very complex economic category. Researchers are usually not sure which variables should be introduced into the model. Often because of different angles, the selected variables are different. As a micro-housing market, the factors are more likely to interact with each other, and the superposition effect is more obvious. It is worth to think which independent variables should be introduced into the model, and which interaction effects between variables cannot be ignored. Lasso can autoselect influencing factors with strong explanatory ability, which can well solve the above problems.

Lasso is a coefficient compression variable selection method proposed by Tibshirani in 1996. Its essence is to add an L1 penalty term, so that it can estimate the independent variable coefficients. Let $\xi=\left(\beta_1,\cdots,\beta_{n_1},\gamma_1,\cdots,\gamma_{n_2},\lambda_1,\cdots,\lambda_{n_3},\eta_1,\cdots,\eta_{n_4},\theta_1,\cdots,\theta_{n_5}\right)^T$, then the semi-logarithmic hedonic model based on lasso regression is:

$$\left(\hat{\alpha},\hat{\xi}\right)=\arg\min_{\xi}\sum_i\left(\begin{array}{l}\ln P_i-\alpha_i-\sum_{j=1}^{n_1}\beta_j \ln x_{ij}-\sum_{m=1}^{n_2}r_m \ln x_{im}+\sum_{k=1}^{n_3}\lambda_k x_{im}\ln x_{ij}\\ -\sum_{p=1}^{n_4}\eta_p x_{im_1}x_{im_2}-\sum_{q=1}^{n_5}\theta_q \ln x_{ij_1}\ln x_{ij_2}-\varepsilon_i\end{array}\right)^2 \tag{6}$$

$$s.t.\sum_j\left|\xi_j\right|<t$$

By reducing the value of, the estimated values of some regression coefficients become smaller or equal to 0, to realize variable selection. However, the essence of the model after variable selection is still a multiple regression model. There is still a situation of ignoring the uncertainty of the model itself. Bayes Model Averaging can full use the information in each "suboptimal model", comprehensively consider the priori information and the information provided by the sample. It evaluates the role of each explanatory variable and makes the analysis more scientific. Madigan and Raftery proved that the model established based on the Bayes Model Averaging method has better fitting effect and prediction accuracy than a single deterministic model, and can better ensure the effectiveness of parameter estimation.

Let Lasso choose l variables, and the corresponding parameter vector is $\delta=\left(\delta_1,\cdots,\delta_v\right)$. Any combination of l variables will generate 2^l models, that is, the model space is $M=\left\{M_1,M_2\cdots M_{2^l}\right\}$. Set M_j as the j^{th} model in the model space, $P(M_j)$ is the prior distribution of the model. Under Observational Data D, the posterior probability of M_j is:

$$P(M_j\left|D\right.)=\frac{P(D\left|M_j\right.)P(M_j)}{\sum_{h=1}^{2^k}P(D\left|M_j\right.)P(M_j)} \tag{7}$$

δ_j is the corresponding parameter vector of M_j, $P(\delta_j\left|M_j\right.)$ is the prior probability distribution of the parameter vector δ_j, $P(D\left|\delta_j,M_j\right.)$ is the corresponding likelihood function of M_j, then the marginal likelihood function of M_j is:

$$P(D|M_j) = \int P(D|\delta_j, M_j)P(\delta_j|M_j)d\delta_j \tag{8}$$

Taking the posterior probability of the model M_j as the weight, the posterior probability of the parameter vector δ can be obtained in the form of weighted average:

$$P(\delta|D) = \sum_{j=1}^{2^l} P(\delta_j|M_j, D)P(M_j|D) \tag{9}$$

The posterior mean and posterior variance of the parameter vector δ can be calculated from formula x-formula x, respectively:

$$E(\delta|D) = \sum_{j=1}^{2^l} E(\delta_j|D, M_j)P(M_j|D) \tag{10}$$

$$Var(\delta|D) = \sum_{j=1}^{2^l} Var(\delta_j|D, M_j)P(M_j|D) \tag{11}$$

Among them, $P(\delta_j|M_j, D)$ is the posterior probability of δ_j under the alternative model M_j, $E(\delta_j|D, M_j)$ is the posterior mean of the parameter vector in the single model, and $Var(\delta_j|D, M_j)$ is the posterior variance of the parameter vector in the single model.

4. An Empirical Test of the Factors Influencing Intra-city House Price Differences

In 2012, Chengdu was included in the list of 50 large and medium-sized cities in the country. The rigid demand for housing in Chengdu has maintained a growth trend (Gao-lu, 2015). The urban economic development indicates anisotropic economic development in all directions, with significant differences in public services and population, which has a deep impact on the pattern of housing price differentiation in Chengdu. Therefore, this paper select Chengdu as the study area.

Chengdu real estate market is divided into four directions: east, west, north and south, with 79 areas. Collect its 82-months data from January 2010 to October 2016, from Chengdu Zhenghe Real Estate Consulting Joint stock company, Anjuke and China real estate market network.

Chengdu's urban economic development and population expand to the west, northwest and southwest, and high housing price agglomeration areas extend from the city center to the west and south. First, this paper selects the Waiguanghua area in the city's west for empirical analysis. The model results are shown in Table 5.

Table 5. Regression Results of the Average Lasso Bayesian Model

Variable	$PIP(\%)$	EV	SD	Variable	$PIP(\%)$	EV	SD
$\ln br$	90.2	0.0983	0.0516	$\ln sa$	27.3	-0.0035	0.0071
ra	7.0	-0.0007	0.0043	$\ln lp$	87.0	0.0693	0.0397
bu	11.4	0.000	0.0005	$\ln lhp$	100.0	0.2839	0.0525
$\ln oia$	100.0	0.0046	0.0013	$\ln pd$	94.5	0.0826	0.0300
$\ln st$	17.4	-0.0060	0.0167	$nbr * \ln lpd$	97.4	0.3910	0.1469

Note. PIP, EV and SD respectively represent the posterior inclusion probability, posterior mean and posterior standard deviation.

Observe Table 5, (1) Location attributes. Location economic attributes have an extremely important positive role in promoting housing prices in the area. Location humanistic and social attributes have an impact, but the explanation degree is relatively low. The specific performance is that the posterior probability of the industrial population ($\ln oia$) in this area reaches 100.0%, the posterior probability of the resident population ($\ln pd$) in this area is 94.5%, and the posterior mean values are all positive; The posterior probability values of subway number (ra) and commercial value (bu) are 7.0% and 11.4% respectively, which are close to the critical value of 10%.(2) Market relations. The market environment has a great positive impact on residential prices in the area, followed by supply and demand. Specifically, the posterior probability value of brand projects ($\ln br$) is 90.2%, the posterior probability value of land prices ($\ln lp$) is 87.0%, and the posterior probability values are all positive. The ($\ln st$) posterior probability value of stock is 17.4%, and the posterior probability value of sales volume ($\ln sa$) is 27.3%, far less than 100.0%. And their posterior mean values are all negative, indicating that the stock volume and sales have a low explanation degree for residential prices, and have a negative impact. (3) Related areas. The residential price in the associated area ($\ln lhp$) has a great positive impact on the residential price in the area, and its posterior probability value is 100.0%. (4) The posterior inclusion probability value of the interaction effect between the new brand projects and the associated residential price ($nbr * \ln lpd$) is 97.4%, and the posterior mean is positive. It indicates the positive interaction effect of the residential price in the associated area and the new brand projects on the residential price in the area.

From the regression coefficients and posterior probabilities of the five optimal models in Table 6, the lasso Bayesian mean model is obtained:

$$\begin{aligned}\ln y = 0.297 * model1 + 0.166 * model2 + 0.079 * model3 \\ + 0.058 * model4 + 0.055 * model5\end{aligned} \tag{12}$$

$$\begin{aligned}\ln y = 3.1736 + 0.0705 * \ln br + 0.0029 * \ln oia + 0.00001 * \ln st \\ - 0.0022 * \ln sa + 0.0445 * \ln lp + 0.1809 * \ln lhp + 0.0523 * \ln pd \\ + 0.2633 * (nbr * \ln lpd)\end{aligned} \tag{13}$$

It can be seen that the interaction effect of the new brand projects and the associated residential price ($nbr*\ln lpd$) has the greatest impact on the housing price in the Outer Guanghua area, with a coefficient of 0.2633. That is, the interaction effect of the new brand projects and the associated residential price ($nbr*\ln lpd$) increases by 1%, the residential price in the Outer Guanghua area increases by 0.2633%. Followed by the associated residential price ($\ln lhp$), brand projects ($\ln br$), and land price ($\ln lp$), the residential price increases by 0.1809%, 0.0705%, and 0.0445%, respectively.

Table 6. Top Five Optimal Single Models

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
$\ln br$	0.098828**	0.09751***	0.10690***	0.17220***	0.11900***
$\ln oia$	0.00473***	0.00425**	0.00437***	0.00382***	0.00391**
$\ln st$	-	-	-0.02478**	-	-0.05947***
$\ln sa$	-	-0.01325***	-	-	-
$\ln lp$	0.07282***	0.07744***	0.07067***	-	0.08003***
$\ln lhp$	0.28560***	0.27760***	0.27250***	0.26550***	0.23880***
$\ln pd$	0.09300***	0.08010***	0.06337***	0.10970***	-
$nbr*\ln lpd$	0.39650***	0.40730***	0.39470***	0.42780***	0.39860***
$cons$	4.51900***	4.80300***	5.16400***	4.85300***	6.26700***
Number of variables	6	7	7	5	6
Adjusted R^2	0.930	0.933	0.931	0.923	0.927
BIC	-186.2	-185.1	-183.6	-183.0	-182.9
posterior probability	0.297	0.166	0.079	0.058	0.055

Note. *, ** and *** respectively indicate significant at 10%, 5% and 1% level of significance.

4.1 Comparative Analysis among Different Areas of the Four Directions

Chengdu is divided into four directions, south, east, north and west. According to the suggestion of Chengdu Zhenghe Real Estate Consulting Joint stock company, from the aspects of residential price, volume, and importance of the area, this paper selects 7 representative areas for each direction. Model results are shown in Table 7.

Observe Table 7, (1) Location attributes. Location economic attributes have a greater impact on residential prices in the eastern and have interactive effects. Location humanistic and social attributes have less impact. The specific manifestations are that the resident population ($\ln pd$) and industrial population ($\ln oia$) in this area appear more frequently, and the interaction effect is significant. The commercial (bu) and subway (ra) occasionally significant. (2) Market relations. The relationship between supply and demand has a greater impact on residential prices in the area, and the market environment has a smaller impact. Specifically, the frequency of the inventory-sales ratio (ssr) is also high, and the sales ($\ln sa$), land price ($\ln lp$), and brand projects(br) only have an impact on some areas. (3) Associated area. The associated residential price ($\ln lhp$) has a great influence on the residential price in each area in the east of the city. Its posterior probability is almost 100.0%, and the posterior mean is positive. (4) There are interactive effects in almost every area in the eastern, but the interaction situations are different. There are differences in the influencing factors of housing prices in the area east of the city. Even if there are the same influencing factors, the mode of action and the size of the contribution are also different.

Table 7. Modeling Results for 7 Areas in the East Direction of the City

	Baili	Beihu	Chenglong Road	Chegnren Road	Sichuan Normal University	Cuijiadian	Damian South
Main influencing factors and their regression coefficients	$\ln pd$ 2.1786*** (20.9)	$\ln lhp$ 0.8075*** (100.0)	$\ln br$ 0.0045*** (39.5)	$\ln lhp$ 0.2933*** (100.0)	ra 0.1940*** (100.0)	$\ln br$ 0.0020*** (82.2)	$\ln br$ 0.1680*** (100.0)
	$\ln lhp * \ln pd$ 0.2321*** (87.9)	ssr 0.0005***(95.6)	bu 0.0011*** (70.6)	$\ln lhp * \ln pd$ -0.0223*** (15.7)	$\ln lp$ 0.0155*** (18.7)	$\ln lhp$ 0.5729*** (100.0)	$\ln lhp$ 0.7982*** (100.0)
	$\ln sa * \ln pd$ 0.0369*** (100.0)	$\ln pd$ -0.0293*** (100.0)	$\ln oia$ 0.0003** (37.7)	$ra * \ln oia$ 0.0027*** (80.5)	$\ln lhp$ 0.4976*** (96.5)	$lssr * \ln oia$ 0.0324*** (100.0)	$\ln oia$ -0.0172*** (42.4)
	$ra * ssr$ 0.0002*** (40.6)	-	$\ln lhp * ssr$ 0.0166*** (100.0)	$\ln nbr * \ln oia$ 0.0026*** (100.0)	$\ln oia * lssr$ 0.0005*** (9.2)	$\ln lhp * ssr$ 0.0062*** (41.5)	$\ln oia * sa$ -0.0244*** (100.0)
	$ra * \ln loia$ 0.0010*** (33.7)	-	$pa * \ln lhp$ 0.0092***(100.0)	-	$\ln nbr * \ln sa$ -0.0154*** (100.0)	$ssr * \ln oia$ 0.0084*** (46.9)	-
	$\ln sa * ssr$ 0.00008*** (40.6)	-	-	-	-	$ra * ssr$ -0.0002*** (82.5)	-

Note. The values below variables are regression coefficients, and values in parentheses are posterior inclusion probabilities (%).

Observe Table 8, (1) Location attributes. Location economic attributes have a great impact on housing prices in the southern part of the city, while location humanistic and social attributes have very little impact. The specific performance is that the resident population ($\ln pd$) and industrial population ($\ln oia$) of this area are significant in almost every area. The public facilities value is almost insignificant, only the education level (sc) is significant in the Financial City area. (2) Market relations. The relationship between supply and demand has a great impact on housing prices in the area, and the market environment has a smaller impact. Specifically, the stock-to-sales ratio (ssr) , sales ($\ln sa$) and stock ($\ln st$) are significantly more frequent. and there may be negative effects. (3) Land price. The associated residential price ($\ln lhp$) has a great influence on the housing price in each area in the south of the city, and there is a strong interaction effect. (4) The influencing factors of housing prices are quite different and complex. The interaction effect exists but is weak.

Table 8. Modeling Results for 7 Areas in the South Direction of the City

	Venture Road	Financial City	Hongpailou	Huafu 1 phase	Huafu 2 phase	Huafu 3 phase	Huayang Old Town
Main influencing factors and their regression coefficients	$\ln lhp$ 0.1522*** (52.1)	sc -0.1603** (100.0)	$\ln lhp$ 0.2262*** (100.0)	$\ln nbr$ 0.0143*** (54.7)	$\ln lhp$ 1.2493*** (100.0)	ssr 0.0005*** (100.0)	$\ln oia$ 0.0007*** (70.5)
	$\ln sa * \ln pd$ 0.0268*** (100.0)	$\ln st$ -0.0283*** (73.8)	$\ln lp$ 0.0294*** (100.0)	$\ln lhp$ 0.5571*** (100.0)	$nbr * \ln lhp$ 0.4696*** (100.0)	$\ln oia$ 0.0004*** (100.0)	$\ln lhp$ 0.2896*** (100.0)
	$pa * \ln lhp$ 0.0140*** (52.1)	$\ln lhp$ 0.6022*** (100.0)	$\ln st$ -0.0556*** (100.0)	$nbr * \ln loia$ 0.0017*** (36.8)	-	$\ln st$ 0.0099*** (51.6)	$\ln lhp * ssr$ 0.0044*** (67.9)
	-	$\ln lhp * ssr$ -0.0035*** (30)	$\ln lhp * \ln oia$ -0.0302*** (33.4)	$nbr * \ln sa$ 0.0018*** (100.0)	-	$\ln lhp$ 0.3254*** (100.0)	-
	-	$ra * \ln loia$ 0.0004*** (100.0)	$nbr * \ln oia$ 0.0097*** (80.8)	-	-	$\ln sa * \ln oia$ 0.0099*** (51.6)	-

Observe Table 9, (1) Location attributes. The economic attributes of the location have a certain impact on the housing price in the southern part of the city, and the humanistic and social attributes have a great impact. The specific performance is that the subway (ra), commercial (bu), and environment (wa) are significant in different areas, and the resident population ($\ln pd$) and industrial population ($\ln oia$) in this area are more frequent. (2) Market relations. The relationship between supply and demand and the market environment has a certain impact. (3) Related areas. Relative housing prices ($\ln lhp$) have a less impact on housing prices in each area in the city's west than in the east. (4) The influencing factors of housing prices in the west of the city are more complex than those in the east and south of the city, and the main influencing factors in each area are quite different.

Table 9. Modeling Results for 7 Areas in the West Direction of the City

	High-tech Western District Start-up Area	High-tech Western District Expansion Area	Guanghua	Guanghua Avenue (Wenjiang)	Hongguang (Pixian)	Huangtianba	Wai Guanghua
Main influencing factors and their regression coefficients	ra 0.1700*** (100.0)	$\ln br$ 0.0010*** (100.0)	wa 0.0278*** (21.5)	$\ln loia$ 0.0122*** (100.0)	$\ln br$ 0.0010*** (52.0)	$lssr$ -0.0028*** (70.5)	$\ln br$ 0.0983*** (90.2)
	$\ln loia$ 0.0061*** (100.0)	$\ln oia$ 0.0004*** (48.7)	$\ln loia$ 0.1069*** (48.1)	$\ln sa$ -0.0278*** (100.0)	$\ln st$ 0.1210*** (100.0)	$\ln loia$ 0.0082*** (67.5)	$\ln lhp$ 0.2837*** (100.0)
	$\ln pd$ -0.0367*** (43.0)	$\ln pd$ -0.0719*** (98.6)	$\ln st$ 0.0635*** (97.8)	$\ln lhp$ 0.2403*** (100.0)	$\ln sa$ 0.0016** (17.1)	$\ln lhp$ 0.3944*** (100.0)	$\ln oia$ 0.0046*** (100.0)
	-	$lssr * \ln oia$ 0.0104*** (58.6)	$\ln lp$ -0.0564*** (26.3)	$\ln lhp * ssr$ -0.0035*** (19.2)	-	$lssr * \ln oia$ -0.0131*** (34.4)	$\ln sa$ -0.0035*** (27.3)
	-	-	$pa * \ln lpd$ 0.0091*** (48.1)	-	-	$ra * \ln loia$ 0.0026*** (57)	$\ln lp$ 0.0694*** (87)
	-	-	$nbr * ra$ 0.0016*** (76.1)	-	-	-	$\ln pd$ 0.0825*** (94.5)
	-	-	-	-	-	-	$nbr * \ln lpd$ 0.3910*** (97.4)

Observing Table 10, (1) Location attributes. Economic attributes, humanistic and social attributes all have a great impact on the housing price in the area. (2) Market relations. The relationship between supply and demand and the market environment has a certain impact. (3) Related area. The related housing price ($\ln lhp$) has a great influence on the housing price in each area in the south of the city. (4) The influencing factors of the housing price in the north of the city are the simplest.

Table 10. Modeling Results for 7 Zones in the Northern Direction of the City

	Banyuan	Chengbei 198	Dafeng	Fenghuang Mountain	Jiulidi	Commercial City	Sima Bridge
Main influencing factors and their regression coefficients	$\ln sa$	$\ln br$	bu	$\ln lhp$	ssr	$\ln lhp$	$\ln br$
	-0.0052***	0.1115***	0.0160***	0.5035***	0.00001**	0.7251***	3.1974***
	(83.1)	(46.9)	(100.0)	(50)	(29.4)	(100.0)	(23.7)
	$\ln lhp$	$\ln lhp$	$\ln oia$	$\ln lhp * \ln pd$	$\ln loia$		$\ln lp$
	0.7550***	0.2195***	0.0004**	0.5035***	0.0055***	-	0.0841***
	(100.0)	(59.4)	(18.2)	(50.0)	(100.0)		(73.3)
	$\ln sa * \ln oia$	$ra * \ln loia$	$\ln lhp$		$\ln pd$		$\ln pd$
	-0.0080***	0.1011***	0.4504***	-	-0.0602***(87.1)	-	0.0833***
	(100.0)	(84.8)	(100.0)				(60.5)
		$pa * \ln lhp$			$lssr * \ln oia$		$\ln lhp * \ln pd$
	-	0.0098***	-	-	0.0015***	-	0.0213***
		(67.2)			(39.5)		(30.1)

5. Conclusion and Recommendations

Currently, the overall excess of China's urban housing market and the shortage of supply in some areas coexist, and the market differentiation is very obvious. As the city is in a period of rapid urbanization development, the humanistic, social and economic attributes, market relations and related areas have an important impact on the development of the housing market. Exploring the internal mechanism of residential price differentiation in the area has important practical significance for the healthy development of the real estate market in the future.

5.1 Conclusion

First, the attributes of the area itself have a positive impact on the housing price. The greater the attribute value of the area, the higher the housing price. Economic attributes have a greatest impact on the housing price of the area, followed by humanistic and social attributes, showing that the unbalanced and insufficient development within the city is an important internal factor that leads to the differentiation of housing prices.

Second, the price of the following pairs of housing in the market has a certain impact, and the impact direction can be positive or negative. The relationship between supply and demand has a greatest impact, followed by the market environment. The imbalance between housing supply and demand is an important internal factor that leads to the differentiation of housing prices in urban areas.

Third, in terms of the impact of the related area, the price of related housing has a significant positive impact, reflecting that the related housing market has a great impact on the relationship between supply and demand and market environment of the residential market in the area. It shows that the mobility of the purchase location and the contagion of housing prices are the main reasons for the differentiation, and there is a strong interaction effect between the influencing factors.

Specifically, the main influencing factor of urban residential prices is the residential price in the associated area ($\ln lhp$). This variable is significant in 92% of the areas, which has a positive spillover effect on housing prices and a spatial proximity effect. The industrial population (oia) and the resident population in the area ($\ln pd$) have a large positive impact on the residential price, and the interaction effect is large, which will aggravate the differentiation pattern of residential prices in the area within the city. Stock ($\ln st$) and sales ($\ln sa$) are likely to have a negative impact on housing prices. Subway (ra), commercial (bu) and other public services have a certain impact on residential prices, with a little impact on the pattern of residential price differences. The public facilities factor in the southwest has a stronger explanatory power of housing prices than other directions. Land prices ($\ln lp$) impact housing prices in the central agglomeration area. There are neglected interactive effects on the residential price indicators in the location, with large differences in the influencing factors of residential prices in various areas within the city.

In view of this, one-size-fits-all regulation is not conducive to the stable development of the real estate market. In regulating housing prices, not only need traditional economic and financial policies but also should focus on the optimization of urban internal policies.

5.2 Suggestion

The following recommendations are given to alleviate the unbalanced development and partial underdevelopment of residential market within city.

First, speculative demand exists in areas where indicators such as housing price, sales, and stock-to-sales ratio in the related area have a significant impact, which shows active relationship between the related market and the residential market in the area. The government should focus on providing price-limited housing, implement purchase restriction policies, reform the real estate tax system on second-hand housing transaction, strengthen the supervision of land-related financing, squeeze the space for speculative investment in housing, prevent expanding local real estate bubbles, and reduce the pressure on rigid demand to purchase houses in central agglomeration areas.

Second, rigid demand for housing exists in areas with large industrial population, resident population, and high demand value. The government should focus on providing affordable housing, increase the construction of them, and supplement market-oriented operation, which can reduce the pressure on housing rents and housing prices in such areas. Focus on the optimization of public facilities, provide basic material conditions for attracting industrial talents, reducing the shortage of housing and easing the pressure on the increasing housing prices.

Third, for areas with low significance of economic attributes and market relationship, optimize the location attributes and enhance the value of the humanistic and social attributes of the area. Improve public facilities and services, enhance the livability of the area, improve the investment and operation environment of enterprises, accelerate upgrading the area’s economy, and further enhance the real estate brand projects.

References

Bangura, M., & Lee, C. (2019). House price diffusion of housing submarkets in Greater Sydney. *Housing Studies*, *35*, 1110-1141. https://doi.org/10.1080/02673037.2019.1648772

Fan, J., & Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association*, *96*, 1348-1360. https://doi.org/10.1198/016214501753382273

Feldkircher, M., Horvath, R., & Rusňák, M. (2013). Exchange Market Pressures During the Financial Crisis: A Bayesian Model Averaging Evidence. *International Finance eJournal*. https://doi.org/10.2139/ssrn.2289755

Gao-lu, Z. (2015). *The Effect of Central Business District on House Prices in Chengdu Metropolitan Area: A Hedonic Approach*. https://doi.org/10.2991/cas-15.2015.83

Goodman, A. C., & Thibodeau, T. G. (1998). Housing market segmentation. *Journal of housing economics*, *7*(2), 121-143. https://doi.org/10.1006/jhec.1998.0229

Guan, J., & Gao, J. (2018). Time and Space Analysis of House Price in Mainland China in the Last 10 Years. *Modern Economy*, *09*, 1520-1532. https://doi.org/10.4236/me.2018.99096

He, X. J. (2018). U*nder the general equilibrium perspective of house price decision model: An empirical research of the effect of the property tax of China*.

Hu, J., Xiong, X., Cai, Y., & Yuan, F. (2020). The Ripple Effect and Spatiotemporal Dynamics of Intra-Urban Housing Prices at the Submarket Level in Shanghai, China. *Sustainability*. https://doi.org/10.3390/su12125073

Lindahl-Kiessling, K., & Landberg, H. (1995). Population, economic development, and the environment. *Population and Development Review*, *21*, 415. https://doi.org/10.2307/2137503

Niu, F., & Liu, W. (2017). Modeling urban housing price: The perspective of household activity demand. *Journal of Geographical Sciences*, *27*, 619-630. https://doi.org/10.1007/s11442-017-1396-2

Panduro, T. E., & Thorsen, B. J. (2014). Evaluating two model reduction approaches for large scale hedonic models sensitive to omitted variables and multicollinearity. *Letters in Spatial and Resource Sciences*, *7*, 85-102. https://doi.org/10.1007/s12076-013-0103-x

Sah, V., Conroy, S. J., & Narwold, A. (2016). Estimating School Proximity Effects on Housing Prices: Importance of Robust Spatial Controls in Hedonic Estimations. *The Journal of Real Estate Finance and Economics*, *53*, 50-76. https://doi.org/10.1007/s11146-015-9520-5

Stirk-Wang, L., & Thorsnes, P. (2017). *Geographic variation in intra-city house price appreciation over the boom-bust cycle: evidence from Auckland, NZ.*

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the royal statistical society series b-methodological*, *58*, 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Wu, J., Deng, Y., & Liu, H. (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

Xin, S. J., & Khalid, K. (2018). Modelling House Price Using Ridge Regression and Lasso Regression. *International Journal of Engineering & Technology*. https://doi.org/10.14419/ijet.v7i4.30.22378

Yang, Z., Fan, Y., & Zhao, L. (2018). A Reexamination of Housing Price and Household Consumption in China: The Dual Role of Housing Consumption and Housing Investment. *The Journal of Real Estate Finance and Economics*, *56*, 472-499. https://doi.org/10.1007/s11146-017-9648-6

Zhu, B., Füss, R., & Rottke, N.B. (2013). Spatial Linkages in Returns and Volatilities Among U.S. Regional Housing Markets. *Wiley-Blackwell: Real Estate Economics*. https://doi.org/10.1111/j.1540-6229.2012.00337.x

Zhu, L., & Zhang, H. (2020). Analysis of the diffusion effect of urban housing prices in China based on the spatial-temporal model. *Cities*, 103015. https://doi.org/10.1016/j.cities.2020.103015

Published by SCHOLINK INC.