Empirical Research on Determinants of Housing Prices in Large Cities

Mingyuan Liu(✉)

Guanghua Cambridge International School, Shanghai 201319, China
Mystery@email.cu.edu.kg

Abstract. This dissertation mainly analyses the determinants of housing prices in large cities. For different cities, reasons behind variations in housing prices are also distinguishable. Therefore, through both literature research and empirical analysis, 5 potential determinants of average prices of commercial housing (floor space of buildings completed, city per capita disposable income, investment in real estate development, end-year permanent household size and GDP per capita) are analysed respectively, mainly by correlational analysis and stepwise regression method. The result reached is that city per capita disposable income is always a factor affecting average price of commercial housing, investment on real estate development is a factor affecting average price of commercial housing in most cases, despite some cities banning investment strongly, and end-year permanent household size is occasionally a factor, largely depending on other factors such as the speculative demand and willingness to purchase houses of local population. Other two possible determinants have no influence on average price of commercial housing.

Keywords: housing price · determinants · large cities

1 Introduction

In recent years, considerable housing price rises on a large scale could be witnessed nearly all over China, especially in cities with relatively large scale such as Shanghai and Beijing. Taking housing price in Shanghai as an example, as data in Dushiquan Website have shown, average housing price in Shanghai has grown from 49616 yuan per meter square in December 2017 to 53991 yuan per meter square in December 2021, reaching a 9% growth in approximately 4 years.

Despite such excess growth in some areas, however, some other areas have witnessed a slower growth and a much smaller value of their housing prices, even with similar geographical position. Zhenjiang City can be taken as an example. Based on data in Anjuke Website, although Zhenjiang and Shanghai all belong to the Eastern region of China, the average housing price in Zhenjiang remained significantly lower than Shanghai, reaching only 9049 yuan per meter square in 2017 and 9709 yuan per meter square in 2021, less than 20% of average housing price in Shanghai in both years. Meanwhile, the growth rate of Zhenjiang average housing prices was only 7.2% from 2017 to 2021,

F. Balli et al. (Eds.): ESFCT 2022, AEBMR 226, pp. 1024–1035, 2022.
https://doi.org/10.2991/978-94-6463-052-7_115

lower than Shanghai average housing price growth 9% for 1.8%. Also, according to Jia, in 2013, average housing prices for megacities including Beijing, Shenzhen and Shanghai reached 19157 yuan per meter square, while smaller provincial capitals like Nanchang and Xining reached only 4000 to 6000 yuan per meter square [1]. It can thus be seen that there have been huge differences in housing prices between different cities.

As a crucial industry with a large scale, housing industry has established complex relationships with many other industries, such as financial industry, construction ingredient industries and so on. Meanwhile, house is classified a necessity for Chinese consumers. Thus, the variations in housing prices will not only exert a tremendous impact on consumer living standard but will also affect the stability of Chinese economy. Since such obvious differences in regional housing prices exist, finding out the reasons behind these variations in housing prices is meaningful to not only the improvement of the welfare of consumers who are bearing a high housing price, but also the normal function of the economy. Currently, there has been no complete conclusion of reasons for changes and fluctuations in China housing prices. Therefore, in this dissertation determinants of housing prices in large cities will be investigated.

There have been massive number of dissertations describing the general trends of worldwide and China housing prices, plus the analysis of reasons behind the variations in housing prices. For research on reasons for changes in housing prices, recent Chinese domestic researchers usually chose several variables, demand side or supply side or both, and construct models to discover the effect on housing prices originated from different aspects. For research about Chinese housing prices, Li chose 8 variables including floor space completed, investment in real estate development, land price, city population, average income level, sales area, GDP and loan balance of financial institutions to analyze, finding that only land price, average income level, city population, sales area, investment in real estate development and loan balance of financial institutions had significant impact on Chinese housing prices [2]. Cao analyzed reasons for China housing price variations using data from 1998 to 2013 through empirical model, finding that China housing price increases with percentage of urban population [3]. Hao summarized and analyzed 13 years' data from 1998 to 2012 using grey correlation theory and factor analysis, finding that supply side factors affected China housing price the most, while average disposable income had greatest influence among all demand side factors [4].

Meanwhile, some other researchers focused on mechanisms of housing price generation and reasons for regional housing price differences. Fan used 5 variables including average disposable income, real estate mortgage loan amount, interest rate of mortgage, international capital inflow and expectations to analyze short term changes in housing prices, finding that consumer expectations of future house prices had a significant impact on housing prices all over China, mortgage interest rate had no influence, while international capital inflow had no effect on China middle and western regions, but had a negative effect on eastern region [5]. Chen used panel data model to analyze the relationship between city value and housing price, reaching a conclusion that average disposable income is the largest determinant of housing price differences between cities [6]. Han and Wang did the analysis of regional difference, urbanization and housing price, compared change in factors of urbanization and housing price, including proportion of urban population out of total population, determined that urbanization rate varies in same direction

with housing price [7]. Zhu used stepwise regression to analyze how the income level, economic fluctuations, land supply, credit amount affect housing prices, concluding that these four factors can significantly influence housing prices, income level and land supply directly and the others indirectly [8]. Qiao applied theoretical qualitative analysis based on data from 1998 to 2009 or 2010 about liquidity of Chinese economy and the synchronous rate between housing prices and CPI, outlining problems in Chinese housing market including collusions and economically inefficient pricing strategies of producers [9]. Geng has used regression models to analyze several reasons affecting housing price, including GDP per head, city constructed areas, changes in non-agricultural population, land price index, city investment in houses, and city residential area, concluding that non-agricultural population and land price index had no significant influence on housing prices, while housing price is proportional to investment in house constructions and city scale [10]. Hou summarized housing industry related data in several aspects from 1998 to 2011 and analyzed supply side and demand side factors affecting housing price, giving an insight of composition of cost of real estate development [11].

Overall, these dissertations have mostly used quantitative empirical analysis, using data with long time spans, analyzing the relativity of reasons for Chinese market price, either analyzing the entire China, or only some regions and cities, and the time span generally used was 10 years or less. This dissertation enlarges the time span covered of the variables to 20 years, which could potentially lead to a more universal result.

2 Methodology

2.1 Data Source

The raw data chosen for subsequent mathematical analysis mainly comes from some regional statistical yearbooks, which was published by regional Bureau of Statistics. Some information about Beijing comes from Beijing Macro Database, published by Beijing government, while some GDP per capita information is calculated by regional GDP divided by permanent population. Some city per capita annual disposable income is calculated by monthly average times twelve. As all the raw data comes from government publishment, these pieces of information are relatively reliable, detailed, complete and accurate.

2.2 Price Mechanism Used in Analysis

Demand and supply price mechanism is used in subsequent analysis. This price mechanism states that prices of commodities, including housing products, are determined by both its demand and supply. When demand of a good rises, insufficient supply is provided to fulfil increased demand, thus price will increase. When supply of a good rises, overwhelming provision of the good will lead to a decrease in its price. In this dissertation, housing price is represented by average commercial housing price (average price of commercial housing).

Table 1. Variables and Units

Variables	Units
end-year permanent household size	tens of thousands of people
average price of commercial housing	yuan
investment in real estate development	hundreds of million yuan
GDP per capita	yuan
city per capita disposable income	yuan
Floor space of buildings completed	tens of thousands meter square

2.3 City Scale

City scale in this dissertation is measured through 3 aspects: city economic level, city geographical size, and city population scale. City economic level can be represented by city gross domestic output (GDP), real estate investment and disposable income, while city geographical size can be represented by amount of constructed commercial housing. City population scale can be represented by residential population number.

2.4 Indicator Chosen and Explanation

In the analysis, while the explained variable is average price of commercial housing, there are 5 explanatory variables in total, which accounts for changes in average price of commercia housing. Based on determinants of city scale, these 5 variables chosen are GDP per capita, floor space of buildings completed, investment in real estate development, end-year permanent household size, and city per capita disposable income, as shown in Table 1. The time span covered is 20 years, from 1999 to 2018.

GDP per capita shows the extent of economic development, housing market size and represents demand for commercial houses, which can potentially influence housing prices.

Meanwhile, end-year permanent household size and city per capita disposable income also potentially affects demand of the houses. Floor space of buildings completed represents supply of houses, which may directly determine the commercial housing prices, and is also an important indicator of city size. Investment in real estate development is also a supply-side factor which affects the production of housing products, thus affect commercial housing prices.

There are 3 cities in total chosen for analysis: Beijing, Shanghai, and Shenzhen. These cities are all classified as large cities based on their population number and economic development.

2.5 Stepwise Regression Model

Stepwise regression model is a mathematical method used to determine relationships between two or more variables. There is an approximate linear relationship between

dependent variable and independent variables, and the impact of changes in independent variable on dependent variable can be quantified.

The stepwise regression model contains the stepwise method used in regression process. This method contains two processes. Firstly, if the introduction of a new data makes an old data less significant, then the old data will be eliminated. Secondly, if the introduction of a new data makes an eliminated data significant, then in order to construct a best representing regression function, the significant data will be reintroduced into the function. Finally, a function with the most significant data is constructed.

Suppose there are k number of factors that may influence dependent variable y, from x_1, x_2... to x_k, then the expression of linear relationship between y and x can be considered as:

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_kx_k + b \tag{1}$$

where b is a random variable, and a is regression coefficient which represents the extent of impact of corresponding dependent variable on independent variable. An independent variable will be more influential if a is larger. If a is positive, then independent variable and dependent variable change in same trend, and vice versa.

After the regression process, some data will be eliminated, and the equation will be considered as:

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_{k-t}x_{k-t} + b \tag{2}$$

where t terms disappear. About the actual meaning of Eqs. 1 and 2, x_1 is taken as an example. When x_1 increases for $1/a_1$, y will increase for 1. Despite constant term a_0 and random variable b, similar process happens on other terms in Eqs. 1 and 2.

Meanwhile, this model has some premises on data used in analysis. Firstly, there must be more than 2 independent variables. Secondly, the dependent variable must be a continuous variable. Thirdly, the independent variables must be independent from each other.

3 Results and Discussion

3.1 Trend of Housing Prices and Other Indicators

After the collection of data, the values of 3 indicators (explanatory variables) out of 5 can be shown in following three graphs in terms of each city. Figure 1 shows average prices of commercial housing of three cities, while Fig. 2 shows GDP per capita, Fig. 3 shows city per capita disposable income.

As the raw data have different digits, in order to put them into one graph to compare, data are processed by dividing power of ten, not losing the tendency of the data while making them more comparable. End-year permanent household size, investment in real estate development and floor space of buildings completed have almost no change for Beijing, Shanghai and Shenzhen, fluctuating around fixed values, therefore they do not appear in graphical forms.

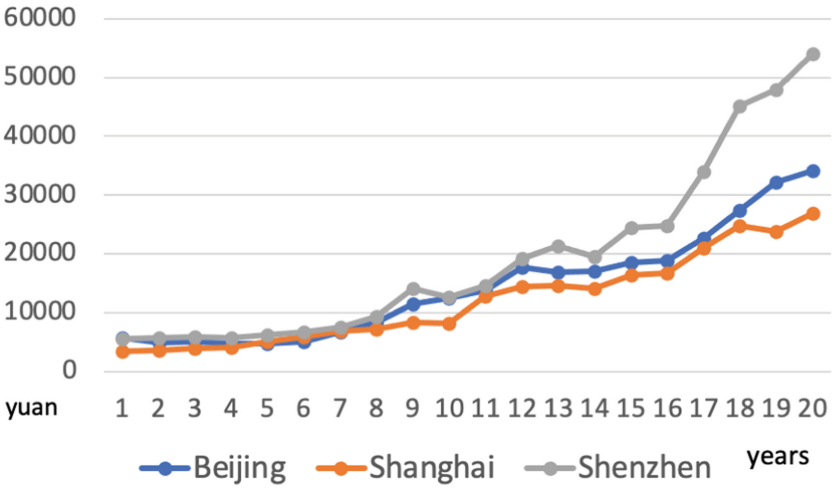

Fig. 1. Average price of commercial housing

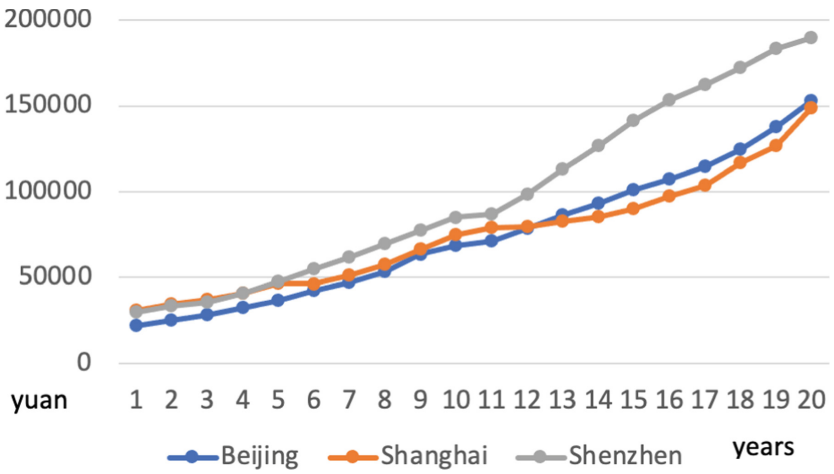

Fig. 2. GDP per capita

In Fig. 1, average price of commercial housing generally increases for 3 cities throughout the whole period, while Shenzhen reaches the highest housing price, followed by Beijing and Shanghai.

In Fig. 2, GDP per capita also rises for 3 cities. Shenzhen also reached highest GDP per capita, while Beijing and Shanghai values are close to each other.

In Fig. 3, Beijing and Shanghai have almost same city per capita disposable income throughout the whole period, while Shenzhen was higher than them at first, yet ending less than Beijing and Shanghai in terms of city per capita disposable income.

Overall, data in three cities share similar increasing trends. In all three figures, GDP per capita and city per capita disposable income can be seen to move with average

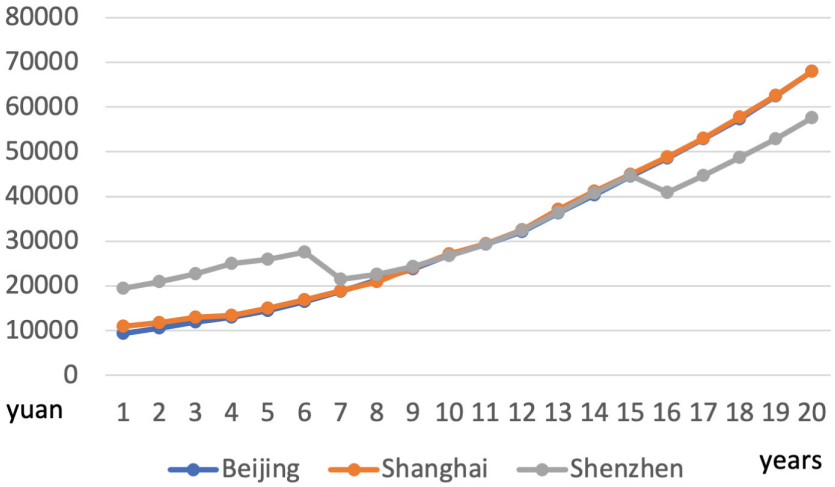

Fig. 3. City per capita disposable income

price of commercial housing in same trend. Thus, GDP per capita and city per capita disposable income is closely related to average price of commercial housing, based on graphs.

This may because that these two factors are demand-side factors, which are easier to change compared to other supply-side factors including floor space of buildings completed which takes several years to change. Therefore, demand-side factor indicators have more similar trends to average price of commercial housing.

Also, average prices of commercial housing continue growing in all three figures, while housing price is generally the highest in Shenzhen, followed by Beijing, then by Shanghai. Housing price growth stopped in three figures in a similar time period, and that may be due to financial crisis in 2008.

3.2 Correlational Analysis

Before the application of regression model, correlational analysis is used first to determine whether explanatory variables have relationship with explained variable, making the result of linear regression model more reliable.

Correlational analysis is done between average price of commercial housing and floor space of buildings completed, city per capita disposable income, investment in real estate development, end-year permanent household size and GDP per capita respectively in terms of Shenzhen data. Pearson correlational coefficient method is used in this process in SPSSAU. By calculating this coefficient, the correlation between dependent variables and independent variables can be found. The calculation process of Pearson coefficient is shown by Eq. 3 below. *E(X)* and *E(Y)* are mean values of independent variable and dependent variable.

$$r = \frac{\sum_{i=1}^{n}(X_i - E(X))(Y_i - E(Y))}{\sqrt{\sum_{i=1}^{n}(X_i - E(X))^2}\sqrt{\sum_{i=1}^{n}(Y_i - E(Y))^2}} \tag{3}$$

Table 2. Correlational Analysis of Shenzhen

Variables	average price of commercial housing
city per capita disposable income	0.954**
end-year permanent household size	0.928**
investment in real estate development	0.965**
floor space of buildings completed	−0.761**
GDP per capita	0.948**

Note: * $p < 0.05$ ** $p < 0.01$

Table 3. Correlational Analysis of Shanghai

Variables	average price of commercial housing
city per capita disposable income	0.988**
end-year permanent household size	0.888**
floor space of buildings completed	−0.327
GDP per capita	0.979**
investment in real estate development	0.981**

Note: * $p < 0.05$ ** $p < 0.01$

The results are shown in Table 2.

It can thus be seen that all 5 explanatory variables are closely related to explained variable. Meanwhile, floor space of buildings completed is negatively related to average price of commercial housing. This is because when floor space of buildings completed increases, supply of houses will increase, which pushes down average price of commercial housing.

Repeat this process on other 2 cities' data, and the results can be shown in Table 3 and Table 4.

In Shanghai city's data, however, the floor space of buildings completed is not related to average housing prices. This is because in Shanghai, the construction area is limited, while most of the available land resources have been occupied by current buildings. Therefore, the availability of addition of floor space of buildings faces constraint, which makes it grow slower than average prices of commercial housing and is unrelated to it.

In terms of Beijing, the floor space of buildings completed is also unrelated to average price of commercial housing. This is because Beijing also has strict rules of construction projects. As the national politics and cultural centre, Beijing weights the environment heavily, thus control the construction project and their quality strictly, and has banned construction projects within Second Ring Road in 2017. Such bans can have negative impact on the amount of new construction projects in Beijing, thus make it grow slower than average price of commercial housing and is not related to average price of commercial housing.

Table 4. Correlational Analysis of Beijing

Variables	average price of commercial housing
city per capita disposable income	0.981**
floor space of buildings completed	−0.027
end-year permanent household size	0.910**
GDP per capita	0.980**
investment in real estate development	0.909**

Note: * p < 0.05 ** p < 0.01

Table 5. Stepwise regression analysis of Shenzhen

Variables	Beta	S.E.	t	p
floor space of buildings completed	−0.125	1.683	−2.163	0.046*
investment in real estate development	0.649	1.462	9.984	0.000**
end-year permanent household size	0.28	6.598	3.363	0.004**

Dependent Variable: average price of commercial housing
Note: *p < 0.05 **p < 0.01

3.3 Empirical Research

In this dissertation, stepwise regression model will be used to find out relationships between the 5 explanatory variables and average price of commercial housing. In the following process, *ince* represents city per capita disposable income, *gdp* represents GDP per capita, *spce* represents floor space of buildings completed, *inv* represents investment in real estate development, *ph* represents end-year permanent household size, and *price* represents average price of commercial housing. Finally, *b* is set as an error term.

Linear regression function can thus be set up as shown below.

$$price = a_0 + a_1 ince + a_2 inv + a_3 gdp + a_4 ph + a_5 spce + b \tag{4}$$

For Shanghai and Beijing city, as the floor space of buildings completed is unrelated to average price of commercial housing, it is eliminated from the regression process of these cities.

The result of stepwise regression of Shenzhen city can be shown in Table 5.

After stepwise regression, three independent variables (floor space of buildings completed, investment in real estate development, and end-year permanent household size) are proved to be the more significant ones influencing dependent variable (average price of commercial housing), two on 0.01 level and 1 on 0.05 level. R square value is 0.981. As a result, 98.1% changes in dependent variable can be explained by three independent variables. Therefore, the model has a strong explanation power. This model passes F test. Therefore, this model is effective. In terms of Shenzhen, the regression equation is:

$$price = -9481.412 - 3.641 * spce + 22.186 * ph + 14.599 * inv \tag{5}$$

Subsequent economic meaning test is conducted. As the coefficient of floor space of buildings completed is negative while other coefficients are positive, floor space of buildings completed and average price of commercial housing change in opposite trend while end-year permanent household size, investment in real estate development and average price of commercial housing change in same trend. This is because floor space of buildings completed represents supply of houses, and price will fall when supply increases. When investment in real estate development increases, there may be speculative demand induced by the expectation that housing prices may fall. The permanent household size also represents demand, and these two demand-side factors can push up prices.

By observing the magnitude of the coefficients, end-year permanent household size influences average price of commercial housing the most, followed by investment in real estate development and floor space of buildings completed in Shenzhen. This may because Shenzhen has a lower population mobility, thus permanent population can exert a greater influence on average price of commercial housing. Average price of commercial housing will increase for 1 yuan if floor space of buildings completed decrease for 1/3.641 m^{vb}. Same 1 yuan increase in average price of commercial housing happens when end-year permanent population size increases for 10000/22.186 people, or when investment in real estate development increases for 100/14.599 million yuan.

Next, the results of stepwise regression of Beijing City are shown in Table 6.

After stepwise regression, two variables (city per capita disposable income and investment in real estate development) are proved to be the most significant factors affecting dependent variable (average price of commercial housing), one on 0.01 level and one on 0.05 level. R square value is 0.971. As a result, 97.1% changes in dependent variable can be explained by two independent variables. Therefore, the model has a strong explanation power. This model passes F test. Thus, this model is effective.

In terms of Beijing, the regression equation is:

$$price = -629.722 + 0.650 * ince - 2.425inv \tag{6}$$

About the economic meaning of Eq. 6, as the coefficient of investment in real estate development is negative while the coefficient of city per capita disposable income is positive, investment in real estate development changes in opposite trend with average price of commercial housing, while city per capita disposable income changes in same trend with average price of commercial housing. The sign of investment in real estate development coefficient is because that in Beijing, the speculative demand induced by investment is less than the increase in supply due to more investment and construction of houses. City per capita disposable income still represents demand and will cause price to rise when it increases and causes demand of the houses to increase.

By observing the magnitude of the coefficients, in Beijing, investment in real estate development has a larger effect on average price of commercial housing than city per capita disposable income. When investment in real estate development decreases for 100/2.425 million yuan, average price of commercial housing will increase for 1 yuan. Similarly, when city per capita income increases for 1/0.650 yuan, average price of commercial housing will also increase for 1 yuan. Permanent population size becomes ineffective here, partially due to most of Beijing permanent population have got dwellings

Table 6. Stepwise regression analysis of Beijing

Variables	Beta	S.E.	t	p
city per capita disposable income	1.292	0.071	9.187	0.000**
investment in real estate development	−0.326	1.047	−2.317	0.033*

Dependent Variable: average price of commercial housing
Note: *p < 0.05 **p < 0.01

Table 7. Stepwise regression analysis of Shanghai

Variables	Beta	S.E.	t	p
city per capita disposable income	0.988	0.015	27.149	0.000**

Dependent Variable: average price of commercial housing
Note: *p < 0.05 **p < 0.01

and have less incentive to purchase for expensive houses, compared to population in Shenzhen, who seek to buy a house after years' work.

Finally, the results of stepwise regression of Shanghai city data are shown in Table 7.

After stepwise regression process, only one variable (city per capita disposable income) is considered as significant to dependent variable (average price of commercial housing) on 0.05 level. R square value is 0.976. As a result, 97.6% of changes in dependent variable can be explained by independent variable. This model passes F test, and the regression equation is:

$$price = -1197.652 + 0.411 * ince \tag{7}$$

When city per capita disposable income increases for 1/0.411 yuan, average price of commercial housing will increase for 1 yuan.

3.4 Summary

Depending on situations of different cities, there are different factors that affect average price of commercial housing. City per capita disposable income is always a factor that affects average price of commercial housing (housing prices), while investment on real estate development affects most cities' housing prices. End-year permanent household size, however, depends on other factors such as permanent population size and mobility of population to have its effect.

4 Conclusion

This dissertation uses mainly stepwise regression to do quantitative analysis. 3 large cities data is used, covering a time span of 20 years from 1999 to 2018, and analysed

the relationship between factors chosen based on large city scale and average price of commercial housing. This dissertation firstly introduces past research on housing prices, concepts of price mechanism, city scale and stepwise model, then used SPSSAU software to do correlational analysis between average price of commercial housing and 5 variables potentially affecting it. After that, 5 variables and the average price of commercial housing are used to build stepwise regression model to reach a result. The result of this dissertation provides some available reference to relative policy making.

References

1. M.L. Jia, Empirical study about the impact of city level on house price changes-Based on 40 major cities in China, 2015, Thesis.
2. H.D. Li, Research on factors influencing Chinese housing prices, 2011, Thesis.
3. W.Z. Cao, Analysis on factors affecting Chinese housing prices, in: BUSINESS, vol. 46, 2015.
4. D.L. Hao, Research on factors affecting Chinese housing prices, 2014, Thesis.
5. L.L. Fan, The Study Of The Influential Factors Of China's Real Estate Price Changes Based On Regional Differences, 2015, Thesis.
6. W.W. Chen, The study of the influence of city value to commercial housing price-empirical research based on 69 large and middle city in China, 2018, Thesis.
7. Z.L. Han, H.W. Wang, Regional difference, urbanization and housing prices-evidence from Chinese housing market, in: Inquiry Into Economic Issues, vol. 2, 2014, pp. 63–70
8. Y.L. Zhu, Econometric research on housing price regional difference in our country, 2012, Thesis.
9. D. Qiao, Research on housing price forming mechanism, 2011, Thesis.
10. B. Geng, Influential factors on Chinese city housing price and spatial difference, 2013, Thesis.
11. C.J. Hou, Research on factors affecting housing price, 2013, Thesis.