

Empirical Analysis of Factors Influencing House Prices in Chinese Provinces

Shuyi Zhang*

School of management, Shanghai University, Shanghai, China

Abstract

With the reform and opening up of China, the country's economic income and purchasing power have been increasing, the most notable feature of which is that the real estate sector accounts for an increasingly high proportion of the total economic output. However, in recent years, housing prices in China's provinces and cities have been increasing at a faster and faster rate, resulting in more and more people being burdened with huge mortgage payments. Therefore, the article will explore the factors influencing the fluctuation of residential real estate prices in 30 provinces, cities and autonomous regions in China and the extent of their impact. Drawing on previous literature and actual circumstances, the article divides the influencing factors into four categories: economic factors, demand factors, supply factors and other factors. A multiple regression model is established with actual data, a unit root test is used to exclude pseudo-regressions, and a Granger analysis is conducted to find out the causal relationship between the explanatory variables and the explained variables. Ultimately, certain recommendations are made for the real estate industry in the context of the current situation in China.

Keywords

Real Estate Price; Empirical Analysis; Co-integration Test; Unit Root Test; Granger Causality Analysis.

1. Introduction

Since ancient times, China has had the traditional concept of "relying on land to move", and in the 21st century, this term has been transformed into "sense of belonging", and whether or not the expression has changed, the essence remains the same: the Chinese people's obsession with buying a home has never waned. From the perspective of economic growth, real estate investment is a channel for the general public to invest in fixed assets, and it is also an important part of the economic development of local governments and the country. Figure 1 below shows the trend of China's GDP and real estate investment amount from 2000-2020, from which it can be found that: in the past 10 years, although both China's GDP and real estate investment amount have maintained growth, the increase has slowed down. According to the relevant data, the annual sales of real estate in 2018 were 149,973 billion CNY, accounting for 6.6% of the total national GDP, analogous to other industries, wholesale and retail trade accounted for 9.4%, construction accounted for 6.9%, and transportation accounted for 4.5%. It is easy to see that the real estate industry also has a strong correlation to economic growth. However, if we look at it from the perspective of social livelihood, in recent years, China's housing prices remain high. Especially in the first-tier cities such as North, Shanghai and Guangzhou, many young people are unable to purchase a property in their lifetime. It is easy to see from this that real estate is closely related to economic development, national happiness and social stability. Therefore, the relevance of exploring the factors that influence the house prices in China is very obvious.

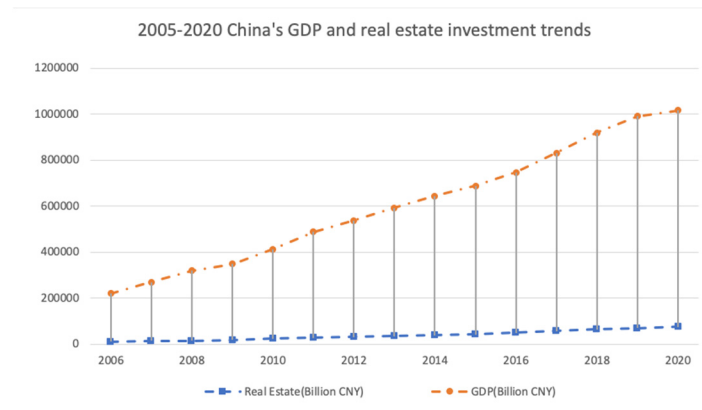


Figure 1. 2005-2020 China's GDP and real estate investment trends

2. Literature Review

From the historical data, there are quite large disparities in real estate prices across Chinese provinces, which has attracted the attention of many scholars. Tan et al[1] conducted a comprehensive study on the factors that cause the fluctuation of house prices. Zheng et al[2], Zhang et al[3], Micheal et al[4], Zhi et al[5], Björn et al[6] and Te et al[7] explored the degree of house price fluctuations from the perspective of macroeconomics and relevant policies introduced by the government; Yue et al[8] and Xing et al[9] explored the influence of regional labor force ratio and aging degree on local house prices from the perspective of population structure. Moreover Xu et al[10] and Hu et al[11] explored the influence of education and the correlation analysis affecting real estate price fluctuations is confirmed from the perspective of residents' consumption behavior.

This paper is divided into six parts: Chapter 1 introduces the development of China's real estate industry; Chapter 2 reviews the relevant literature on real estate influencing factors in China; Chapter 3 identifies the influencing factors; Chapter 4 clarifies the data sources of the article and establishes the corresponding multiple regression model; Chapter 5 is the empirical analysis part; Chapter 6 will combine the analysis results in the paper and China's national conditions to put forward relevant suggestions.

3. Main Influencing Factors

As the real estate sector in China currently covers a wide range of categories, including: commercial office buildings, hotels and guest houses, agricultural land and so on. The differences between the different types of real estate in terms of use and market demand increase. If real estate is analysed as a unified category without differentiating between these categories, it will result in unclear positioning of the research object, which will affect the final results. The residential category is the most relevant as it is the most immediate need of the general public. Therefore, this paper will focus on residential real estate as the main subject of study.

Combining the actual situation and previous related literature, it can be found that the factors affecting real estate prices can be divided into four main areas: macroeconomic factors, demand factors, supply factors and other factors. In the following, each of these four aspects will be analysed in turn.

3.1. Macroeconomic Factors

Most of the indicators related to economic factors are related to the overall economic level of the region. Common indicators are: GDP (Gross Domestic Product), GNP (gross national product), the Engel coefficient, the value of the tertiary industry, CPI (Consumer Price Index)

and so on. Considering the relationship between the various types of index and real estate, we finally selected economic factors: the end of the total storage GDP, residents.

3.2. Demand Factors

Taking into account the minds of most people, the population is an important factor driving the real estate market, real estate sales at the end of the total area of real estate can reflect local consumer market size and the degree of recognition for the real estate industry in the region. The final selected demand factors in this article are: population and housing sales area.

3.3. Supply Factors

The standard for measuring the supply capacity of real estate in a region defaults to the size of the real estate market in the region, and this indicator is usually closely related to the completed area of real estate; in addition, from the perspective of upstream real estate supply, land prices, construction and installation prices, and building materials prices And factors such as the area of land development should also be taken into consideration. Since this article focuses on studying the price of residential real estate from the perspective of the real estate market, the final supply factors selected include: the completed area of real estate.

3.4. Others

In addition to the above three factors that will affect residential prices, there are some additional factors that will also play a role in housing prices, including: the degree of internationalization, the amount of infrastructure investment, and the total mileage of subway construction. However, considering that inflation has the greatest impact on our daily lives. Therefore, the inflation rate is included as an additional consideration in the article.

4. Data and Indicators

This article uses data related to real estate from 30 provinces, municipalities and autonomous regions in China from 2005 to 2017, establishes a multiple regression model, and conducts an empirical analysis of the results to explore the influencing factors and the degree of influence behind China's real estate prices.

4.1. Index Selection

(1) Explained variable: annual real estate prices in various provinces, municipalities and autonomous regions across the country

(2) The explanatory variables

1) Economic factors: GDP, the total amount of residents' storage at the end of the year.

2) Demand factors: population, sales area of houses.

3) Supply factors: real estate completed area.

4) Other factors: inflation rate.

4.2. Data Sources

Most of the data in the article comes from the 2005-2017 China City Statistical Yearbook", "China Real Estate Statistical Yearbook" and the official website of the National Bureau of Statistics. Since the time span spans 12 years, more than 30 provinces are considered, and the amount of data is relatively large. therefore. The following takes Beijing as an example to show some indicator data. The specific figure is shown in Table 1 below:

Data requires some pretreatment before they can carry out correlation and regression analysis. Combining operability and the characteristics of this article, a logarithmic preprocessing method will be used. The most significant advantage of this method: it can eliminate the heteroscedasticity of time series data without changing the long-term stable relationship

between the original variables. All subsequent variables in the article will be analyzed using new logarithmic transformation data.

Table 1. Statistical indicators

Year	House Price (CNY/ m^2)	GDP (Billion CNY)	Housing sales area (m^2)	Resident's total storage at the end of the year	Population (million)	Completed residential area (m^2)	Inflation rate(%)
2005	4456	6969.52	17710533	19620321	1163	20807480	1.825
2006	7375	8117.78	22050296	21573093	1181	21933179	1.47
2007	6162	9846.81	28236529	71543045	1198	4691804	2.5
2008	12418	11115	8227326	74775889	1213	55382293	8.2
2009	13799	12163.03	14704184	86984521	1299.85	16132296	8.5
2010	17782	14113.58	9832771	91134935	1246	14984755	9.1
2011	16852	16251.93	1722616	126490365	1258	13161275	5.1
2012	17022	17879.4	14833746	145658843	1278	15227151	6.1
2013	18553	19800.81	4737721	214045510	1297.5	16920382	5.9
2014	18833	21330.83	11365281	229043566	1316.3	18043430	4.9
2015	22633	23014.59	7956685	239139670	1333	13782203	6.4
2016	28489	25669.13	8345124	281750343	1345	23699489	4.6
2017	32140	28014.94	2277513	3332497723	1363	6040373	1.3

5. Model and Real Political Analysis

This article is based on STATA15.0 software for regression analysis parameter estimation and model testing.

5.1. Descriptive Statistics

Table 2. Descriptive statistics

	LnHousingPrice	Lnpop	Lnsale	Inflation	LnGDP	LnCompleted	LnDeposite
Mean	8.395	7.907	16.242	3.617	9.311	16.427	17.407
Median	8.392	8.175	16.296	4.000	9.457	16.631	17.651
Maximum	10.378	9.327	18.293	5.130	11.404	19.230	21.927
Minimum	6.982	5.267	13.213	1.470	5.267	10.270	6.267
Std.Dev.	0.587	1.0721	0.986	1.062	1.072	1.346	1.637
Skewness	0.369	-1.005	-0.347	-0.730	-0.656	-1.365	-1.809
Kurtosis	3.680	3.143	2.900	2.461	3.491	5.866	12.005
Jarque-bera	15.850	66.02	7.973	39.34	31.89	254.6	1531
Probability	0.000360	0.000	0.0186	0.000	0.000	0.000	0.000
Observations	389	389	389	389	389	389	389

The summary statistics are shown in Table 2. In addition to the basic statistical data such as the mean, minimum, and maximum values of each variable, it is found from the JB test: Because the variable rejects the null hypothesis of normality, the panel data composed of 389 observations in this study all follow normal distribution.

5.2. Correlation Analysis

On the basis of statistical analysis, this article also conducts correlation matrix analysis, the main purpose is to explore the correlation between various variables. According to the data from 2005 to 2017, it can be found that most of the above explanatory variables are positively correlated with housing prices, among which the relationship between GDP and housing prices is relatively close, which means that the higher the GDP of a region, the higher the corresponding housing prices. , This is consistent with our cognition; but among the explanatory variables, there is a significant negative correlation between the population of the province and the housing price, which means that the higher the population, the lower the housing price. This reason may be related to the existence of several populous provinces with less developed economies.

Table 3. Correlation matrix analysis

	LnHousingPrice	Lnpop	Lnsale	Inflation	LnGDP	LnCompleted	LnDeposite
LnHousingPrice	1.000						
Lnpop	-0.056	1.000					
Lnsale	0.194***	0.547***	1.000				
Inflation	0.358***	-0.000	0.209***	1.000			
LnGDP	0.534***	0.643***	0.626***	0.258***	1.000		
LnCompleted	0.377***	0.384***	0.500***	0.553***	0.602***	1.000	
LnDeposite	0.299***	0.396***	0.273***	0.246***	0.451***	0.537***	1.000
	0.000	0.000	0.000	0.000	0.000	0.000	0.000

5.3. Unit Root Test

As the text of the original data for the panel data, it may be a "false regression" phenomenon, hence the need for the unit root test before parameter estimation. There are many unit root inspection methods. In order to ensure the accuracy and reliability of the results, IPS and LLC methods are selected for comparative inspection this time. The specific test results are shown in Table 4.

Can be found from Table 4: P-value and the IPS method LLC were less than 0.05%, and therefore reject the null hypothesis that: prosequence unit root does not exist. Therefore, there is no false regression, and the regression calculation of the model can be carried out.

Table 4. Unit root test

Variable	IPS	IPS(prob)	LLC	LLC(prob)	conclusion
LnHousingPrice	-2.7580	0.029	-12.823	0.027	Stable
Lnpop	-5.9710	0.000	-4.2549	0.000	Stable
Lnsale	-2.6477	0.004	-15.3365	0.001	Stable
Inflation	-3.2624	0.006	-5.8322	0.000	Stable
LnGDP	-9.1864	0.000	-12.5005	0.0231	Stable
LnCompleted	-8.3321	0.000	-13.8409	0.000	Stable
LnDeposit	-7.6576	0.000	-12.9071	0.000	Stable

5.4. Regression Model

First, we need to establish a regression equation about housing prices, as shown below:

$$\text{LnHousingPrice}_{i,t} = C + \beta_1 \text{lnpop}_{i,t} + \beta_2 \text{lnsale}_{i,t} + \beta_3 \text{Inflation}_{i,t} + \beta_4 \text{lnGDP}_{i,t} + \beta_5 \text{lnCompleted}_{i,t} + \beta_6 \text{lnDeposit}_{i,t}, \tag{1}$$

where C is the regression constant; β_i is the regression coefficient, which represents the regression coefficient of city i , $i = 1, 2, \dots, 30$; t represents time, ranging from 2005-2017.

Performing least square regression (OLS) on equation(1) can get:

$$\text{LnHousingPrice} = 5.997(0.000) - 0.3587 \text{lnpop}(0.000) - 0.44037 \text{lnsale}(0.103) + 0.654 \text{Inflation}(0.005) + 0.5367 \text{lnGDP}(0.000) - 0.159 \text{lnCompleted}(0.493) + 0.563 \text{lnDeposit}(0.000) \tag{2}$$

Where $R^2 = 0.5915$, $DW = 0.846$, $F = 89.33$. The values in parentheses represents the probability P corresponding to the value of t.

According to the probability P corresponding to the t value, it can be concluded that the model established in this paper satisfies the significance test. It can be found that the regression equation: R^2 value of 0.5915 model fitting goodness higher, which also proves more appropriate to build the model. At the same time, the variables lnGDP, lnpop, and lnDeposit have coefficients of influence on housing prices: 0.281, 0.296, 0.341, respectively. Description GDP, population and population for the end of the year deposit rate has a significant positive stimulus effect; contrary ln sale, lnCompleted for the price impact coefficient: -0.011 and -0.07, indicating that housing sales area and completion area for the price reverse incentive effect. Specifically: for every 1 percentage point increase in the completed LLC area of houses, house prices will fall by 0.011 percentage points.

5.5. Causal Effects

In econometrics, after constructing a regression model, it is often necessary to make judgments about the causal relationship between economic variables. In actual situations, Granger causality test is used more frequently. Its main idea is to test whether the historical information of an economic variable can be used to predict the future changes of another economic variable. It is not difficult to find out that the important value of Granger test lies in prediction.

In order to further measure the causal relationship between the explanatory variable and the explained variable, this paper introduces Granger causality analysis method to test, the specific situation is shown in Table 5:

Table 5. Granger cointegratance

NULL HYPOTHESIS	Lag period options	F-Statistic	Prob.
Lnpop does not Grange Cause LnHousingPrice	1	0.7303	0.4652
LnHousingPrice does not Grange Cause Lnpop	1	20.8907	0.0000
Lnsale does not Grange Cause LnHousingPrice	1	0.5356	0.5922
LnHousingPrice does not Grange Cause Lnsale	1	8.5098	0.0000
Inflation does not Grange Cause LnHousingPrice	1	-1.6865	0.0917
LnHousingPrice does not Grange Cause Inflation	1	-2.9958	0.0027
LnGDP does not Grange Cause LnHousingPrice	1	28.9632	0.0000
LnHousingPrice does not Grange Cause LnGDP	1	82.1416	0.0000
Lncompleted does not Grange Cause LnHousingPrice	1	6.4993	0.0000
LnHousingPrice does not Grange Cause Lncompleted	1	2.9538	0.0031
LnDeposit does not Grange Cause LnHousingPrice	1	3.1026	0.0019
LnHousingPrice does not Grange Cause LnDeposit	1	-0.8280	0.4077

Since there are many explanatory variables in the model, the economic significance represented by different explanatory variables is quite different, so the choice of lag period is also different. This paper uses the AIC criterion to determine the optimal order as 1. From the results of Granger causality analysis, it can be found that most variables and the explained variable housing prices are not two-way causal relationships, and only a few variables are in a two-way relationship, such as the completed area of houses and GDP. It can be found that when the GDP or the completed area of the house changes, the house price will definitely change, and vice versa. The increase or decrease of house price will affect the change of GDP to some extent. For some one-way influencing factors, such as population. The number of urban population does not necessarily affect local housing prices. According to census data, Chongqing has a permanent population of 30,174,200, ranking first in the country, but housing prices are not the highest in the country. On the contrary, the increase in housing prices will affect the urban population to a certain extent. In recent years, the younger generation has been increasingly calling for fleeing first-tier cities due to high housing prices, resulting in a decline in the number of permanent residents in first-tier cities. This social phenomenon has also been positively verified: there is no mutual causal relationship between housing prices and population.

6. Conclusion

According to the regression model, it can be found that most of the influencing factors have a positive effect on housing prices, and the real estate sales area in the demand factor, the completed area of the real estate in the supply factor, and the inflation rate in other factors are negatively correlated with the housing price; At the same time, these factors and housing prices are not all two-way causal relationships; this can also be found from the coefficients of the regression model: Although these influencing factors have an impact on housing prices, the ability of different factors to affect housing prices also exists different. Based on the results of the model and empirical analysis, we can find:

(1) There is a strong correlation between the real estate industry and economic growth in various regions, and a long-term stable equilibrium relationship is maintained. Therefore, it can be concluded that maintaining stable economic development has a positive and positive effect on stabilizing national real estate prices.

(2) Different regions need to "adjust measures to local conditions", and different cities need to formulate different policies. Take Beijing as an example. In a city with a developed economy and a stable development, housing prices need to be stable or even to restrain housing prices from rising too fast. As Xi Jinping General Bookstore said: houses are for living, not for speculation; Third-tier cities can encourage local residents, even migrants, and foreign businessmen to invest in real estate to boost the local economy.

References

- [1] R. H. Tan, Q. S. He, K. H. Zhou, et al. Administrative hierarchy, housing market inequality, and multilevel determinants: A cross-level analysis of housing prices in China, *Journal of Housing and the Built Environment*, Vol. 34 (2019) No. 3, p.845-868.
- [2] S. G. Zheng and L. Y. A study on house price volatility, estimation of the stance of regulation policies and their impact effects: Empirical analysis based on data from 1998-2014 (*Journal of Finance and Economics, China* 2016), p.98-109.
- [3] H. Zhang, Z. F. Li and B. J. Deng. Policy Uncertainty, Macro Shocks and House Price Volatility: Empirical analysis based on LSTVAR model (*Journal of Financial Research, China* 2015), p. 32-47.
- [4] F. Michael, L.-L. Danilo and T. Andrew. Mapping China's time-varying house price landscape, *Regional Science and Urban Economics*, Vol. 78 (2019), p. 103464.
- [5] T. H. Zhi, Z. F. Li, Z. Q. Jiang, et al. Is there a housing bubble in China? *Emerging Markets Review*, Vol. 39 (2019), p. 120-132.
- [6] R. Björn, S. Moritz and S. Ilhyock. The costs of macroprudential policy, *Journal of International Economics*, Vol. 118 (2019), p.263-282.
- [7] B. Te and H. Cars. When speculators meet suppliers: Positive versus negative feedback in experimental housing markets, *Journal of Economic Dynamics and Control*, Vol. 107 (2019), p. 10 3730.
- [8] H. Yue and X. Y. Zhu. The influence of urban built environment on residential burglary in China: Testing the encounter and enclosure hypotheses, *Criminology & Criminal Justice*, Vol. 21 (2021), p. 508-528.
- [9] L. J. Xing, Y. F. Liu, B. S. Wang et al. An environmental justice study on spatial access to parks for youth by using an improved 2SFCA method in Wuhan, China, *Cities* Vol.96 (2020), p. 102405.
- [10] X. Xu, Y. W. Yu, Y. Xu, et al. A multi-agent reinforcement learning-based data-driven method for home energy management, *IEEE Transactions on Smart Grid*, Vol.11(2020), p. 3201-3211.
- [11] L. R. Hu, S. J. He, Y. Luo et al. A social-media-based approach to assessing the effectiveness of equitable housing policy in mitigating education accessibility induced social inequalities in Shanghai, China, *Land Use Policy*, Vol. 94(2020), p. 104513.