Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/49/20230502

Analysis of Regional Differences and Influencing Factors of Real Estate Prices: An Empirical Study Based on 31 Provinces in China

Liu Xiaodi[1,a,*]

[1]*Nankai University, Haitang Street, Tianjin, China Department of Finance*
a. 2012170@mail.nankai.edu.cn
**corresponding author*

Abstract: Currently, the continuous rise in housing prices in China has brought about social issues regarding housing fairness. Therefore, it is of practical significance to study the influencing factors of China's real estate prices and the effects of these factors. This article first identifies ten variables that have a significant impact on housing prices through analysis, and then conducts empirical analysis separately for the eastern and central-western regions. Based on the correlation and model multicollinearity tests, two corresponding models are established. Subsequently, this article uses a multiple linear regression model for regression and uses a test set to verify the model's effectiveness, conducting quantitative analysis. Combining the analysis results, the factors that have a significant impact on housing prices in the eastern region of China are identified as the sales area of commercial housing, rental prices, and unemployment rates. In the central-western region, the factors with a significant impact on housing prices are disposable income and real estate development costs. Based on these conclusions, this article proposes relevant policy suggestions, which are expected to provide a basis for differentiated macro-control policies by the government.

Keywords: real estate prices, regional differences, impact effect, multiple linear regression

1. Introduction

"House reform" in 1998, as a milestone event in the history of China's reform, ended the welfare housing system that had been in place for nearly fifty years, leading China's real estate industry to truly enter the market, and effectively improving the living quality of the people, laying the foundation for China's high economic growth over the past twenty years. However, while the development of the real estate market has been remarkable, the phenomenon of a five-fold increase in housing prices in China over the past twenty years has also continuously exposed issues such as the need to guarantee basic housing needs and housing fairness [1]. Under this situation, China has successively introduced and implemented a series of real estate regulation policies, aiming to alleviate the heat in the real estate market and maintain its stability by increasing housing supply, imposing purchase restrictions and price controls, and strengthening approval and supervision.

As a fundamental and pillar industry of China's national economy, the real estate industry serves as the "barometer" of China's economic prosperity. Housing prices not only affect the healthy and

DOI: 10.54254/2754-1169/49/20230502

stable development of the real estate market and economic prosperity, but also impact the quality of life and sense of well-being of residents. Therefore, maintaining long-term stability in housing prices is an important goal of macroeconomic regulation. This article aims to explore the key influencing factors of commodity housing prices in recent years based on the current development of the real estate industry in China, and differentiate the regional differences between the eastern and central-western regions to compare and analyze the degree of influence of different factors, ultimately providing relevant basis for the government's differentiated macroeconomic regulation policies.

2. Factors Influencing Real Estate Prices

2.1. Macroeconomic Factors

GDP, as a core indicator of national economic accounting, can effectively measure the overall economic situation of a country; correspondingly, regional GDP can reflect the economic situation of different areas. For the relationship between GDP and housing prices, on one hand, the output of the real estate industry is a component of GDP; on the other hand, the impact of GDP on housing prices is reflected in its representation of the level of regional economic development, which is directly proportional to the purchasing power of residents and the level of effective demand. Therefore, it is expected that housing prices and GDP are positively correlated.

CPI, as a major macroeconomic indicator, can reflect the price level of consumer goods and services purchased by residents. For the relationship between CPI and housing prices, on one hand, whether consumers rent or buy houses, housing is a major part of daily expenditure. Currently, CPI includes rent, which accounts for 13% of the weight system. On the other hand, overall changes in the general price level will lead to changes in housing prices through changes in construction material costs, labor costs and so on. Therefore, it is expected that housing prices and CPI are positively correlated.

The unemployment rate, as a major macroeconomic indicator, can reflect the prosperity or decline of a country's economic situation. For the relationship between the unemployment rate and housing prices, the rise in the unemployment rate reflects the crisis of economic recession. The increase in the unemployed population will lead to a decrease in residents' purchasing power, undermine consumer confidence, and reduce consumers' recognition of housing prices, leading to a delay in the willingness of homebuyers to purchase. Therefore, it is expected that housing prices and the unemployment rate are negatively correlated.

Since most homebuyers choose to purchase through loans, and it involves a high leverage ratio, the housing loan interest rate is an important policy indicator for the country's real estate regulation. By adjusting the housing loan interest rate up or down, the cost of purchasing houses for residents can be influenced, achieving the effects of cooling down or heating up the real estate market respectively. Therefore, according to the principle that interest rates are inversely related to asset prices, it is expected that housing prices and housing loan interest rates are negatively correlated. Wang Tongxu summarized the dynamic adjustment mechanism of the first-home loan interest rates in many cities of China, and believed that the favorable policy of "double reduction" of down payment and loan interest rates is conducive to boosting housing consumption [2].

2.2. Supply Factors

Land supply is the foundation of real estate market development, and the quantity and structure of land supply directly determine the quantity and structure of real estate supply. Land acquisition cost, as the cost paid by real estate enterprises to obtain land use rights, is the most important cost in real estate development, accounting for 30% to 50% of the total cost, and it tends to increase as land

DOI: 10.54254/2754-1169/49/20230502

resources become scarce. Therefore, land acquisition cost is an important indicator from the perspective of real estate supply. Since real estate enterprises will pass on costs to consumers through housing prices, it is expected that housing prices and land acquisition costs are positively correlated. Zhang Ying and Gui Jinsai studied the consistency of land prices, the growth rate of housing prices, and the average profit rate of the real estate industry, and concluded that land prices are the main driver of high housing prices in China [3].

Real estate investment, as an indicator of the total cost of real estate enterprises, is closely related to housing prices, but it is more stable than housing prices because real estate projects generally continue construction once started. For the relationship between real estate investment and housing prices, on one hand, rising housing prices stimulate real estate investment; on the other hand, real estate investment, as a cost for enterprises, will be passed on to consumers through housing prices, so it is expected that housing prices and real estate investment are positively correlated.

Sales area of commercial housing, as a factor with limited elasticity at the supply side, is a direct indicator that consumers can directly contact when buying a house, and it has a more direct correspondence with housing prices and can be explained by supply and demand relationship. Therefore, it is expected that housing prices and sales area of commercial housing are negatively correlated.

Due to the important characteristic of immobility of real estate, its geographic location plays a decisive role in price. In general classification studies, China is usually divided into eastern, central, and western regions or first, second, and third-tier cities. Yi Ying and Liu Meiling analyzed the influencing factors of the price of commercial housing from the perspective of regional differences, and empirically concluded that the price of real estate in eastern China is more sensitive to changes compared to the central and western regions [4]. This article plans to divide the 31 provinces of China into eastern and central-western regions for research, and explore the impact of various factors on these two types of regions.

2.3. Demand Factors

Disposable income, as an important indicator at the demand side, measures the consumer's purchasing power. Since most home buyers use mortgage loans, the repayment amount is limited by disposable income. Therefore, it is expected that housing prices and disposable income are positively correlated.

Renting and buying a house are two alternative residential options for residents and have a mutually substitutable and restrictive relationship. When considering these two options, consumers usually consider renting prices and housing prices as a comparison. Renting prices, as the usage price of houses, can indicate the real demand in the market. Yu Xiaodong compared the costs of renting and buying houses and believed that the renting price of a house usually reflects the real supply and demand relationship in the housing market, while the housing price, due to the large presence of investment purchases and speculation, may deviate from the real supply and demand relationship, thus analyzing the bubble level of housing prices in 27 sample cities across the country [5].

Population density reflects the quantity of housing demand and has a close impact on housing prices, especially in the long term. Therefore, it is expected that housing prices and population density are positively correlated. Xu Huijie studied the impact of population density on housing prices in new first-tier cities based on population migration theory and found a positive correlation between the two, and analyzed the significance of the impact [6].

DOI: 10.54254/2754-1169/49/20230502

3. Data and Methods

This article utilizes panel data related to the prices of residential properties in 31 provinces in China from 2012 to 2021. The data frequency is annual, and the provinces exclude the three regions of Hong Kong, Macau, and Taiwan. The data sources are the National Bureau of Statistics and the People's Bank of China. Based on the factors affecting real estate prices discussed in Chapter 1, the selection of explanatory variables for the model can be made, followed by data collection. The basic characteristics of the data are shown in Table 1, with a total sample size of 310. The data is preprocessed by standardizing all the data for future comparison, and the sample is randomly divided into a training set and a test set in an 8:2 ratio.Descriptive statistics are applied to the housing price data, revealing that the average price, median price, and maximum price of the eastern region (12905.91, 10316.00, and 40526.00 CNY/m^2, respectively) are significantly higher than those of the central-western regions (5834.55, 5558.50, and 9828.00 CNY/m^2, respectively), indicating a noticeable difference.

Table 1: Basic information of data.

Category	Variable Name (Symbol)	Measurement Indicator
Dependent variable	House Price (Y)	Average Sale Price of Houses
Independent variable	GDP (X_1)	Gross Domestic Product
	CPI (X_2)	Consumer Price Index
	Unemployment Rate (X_3)	Urban Registered Unemployment Rate
	Mortgage Interest Rate (X_4)	China: Weighted Average Interest Rate of RMB Loans by Financial Institutions: Personal Housing Loans
	Real Estate Development Cost (X_5)	Land Acquisition Cost
	Real Estate Investment (X_6)	Total Investment Amount Completed this Year
	Area of Commercial Housing (X_7)	Sales Area of Commercial Housing
	Disposable Income (X_8)	Per Capita Disposable Income of Urban Residents
	Rental Price (X_9)	China: CPI: Rental Housing
	Population Density (X_{10})	Urban Population Density

After the initial setup of variables, variable selection is performed based on the correlation coefficient and the Variance Inflation Factor (VIF) to obtain the model used for analysis. According to data from the National Development and Reform Commission, the eastern region of China includes 11 provinces (autonomous regions and municipalities) such as Beijing, while the central-western regions include 20 provinces (autonomous regions and municipalities) such as Shanxi and Sichuan. The variable selection process is as follows: firstly, the correlation coefficient between each explanatory variable and the dependent variable is calculated, and the explanatory variables with a correlation coefficient $|r| \geqslant 0.3$ are retained, while those with $|r| < 0.3$ are removed; then, using a criterion of VIF $\leqslant$ 10, the explanatory variable with the highest VIF is successively removed in each round. This process is repeated until the stopping criterion is met. The purpose of variable selection is to prevent the inclusion of weakly correlated explanatory variables and severe multicollinearity, which would impact the stability and explanatory power of the model.

DOI: 10.54254/2754-1169/49/20230502

For the model setup in the eastern region, the correlation coefficient test retains four explanatory variables: unemployment rate, sales area of residential properties, disposable income, and rental prices, all of which have $|r| \geqslant 0.3$. In the VIF test, after one round of selection (reducing VIF from 11.765 to 1.852), the explanatory variable of disposable income is removed, leaving three variables: unemployment rate, sales area of residential properties, and rental prices. For the model setup in the central-western regions, the correlation coefficient test retains five explanatory variables: GDP, real estate development costs, real estate investment, disposable income, and rental prices, all of which have $|r| \geqslant 0.3$. In the VIF test, the model meets the criterion in the first round (VIF = 3.205), so all five variables are retained. The model used in this study is a multiple linear regression model, and the OLS regression method is employed. The basic form of the model is as follows: $Y = \beta_0 + \Sigma_i \beta_i X_i + \mu$. The subsequent research process involves analyzing the regression model results for the eastern and central-western regions separately, conducting tests using the test set, and then presenting the analysis results, along with recommendations and insights based on the findings.

4. Results and Discussion

4.1. Eastern Region

The coefficients of each explanatory variable in the model are shown in Table 2. For the training set results, the intercept term of the model is -0.0272 (-0.340), and the goodness of fit is 0.460. The unemployment rate, commercial housing sales area, and rental prices are significant at the 1% significance level. It can be seen that macroeconomic indicators and supply-demand indicators are closely related to eastern housing prices. The signs of the coefficients of the explanatory variables are consistent with expectations. The coefficients of the three variables are similar in magnitude and significance, indicating that the unemployment rate, commercial housing sales area, and rental prices have a similar impact on housing prices. The model is tested using the test set, and the intercept term of the model is 0.1214 (0.562), while the goodness of fit is 0.357, slightly lower. The significance of the unemployment rate is slightly lower, possibly due to the reduction in sample size. The signs and sizes of the coefficients of the variables are consistent with those of the training set, indicating that the model is reasonable.

Table 2: Regression results of the eastern region model.

	Unemployment rate	Commodity house sales area	Rental prices
beta (training set)	-0.3930**	-0.4385**	-0.4090**
beta (test set)	-0.3922*	-0.4284**	-0.3960**

Note: *、**、*** represent significance at the 5%, 1%, 0.1% level, respectively. The same applies to the following table.

4.2. Central-Western Region

The coefficients of each explanatory variable in the model are shown in Table 3. For the training set results, the intercept term of the model is 0.3272** (7.149), and the goodness of fit is 0.688. The real estate development cost and disposable income are significant at the 1% significance level, while GDP, real estate investment, and rental prices are not significant at the 5% significance level. It can be seen that supply-demand indicators are closely related to central-western housing prices. The signs of the coefficients of the two significant variables are consistent with expectations, and the coefficient of disposable income is higher and more significant, indicating that this factor has

DOI: 10.54254/2754-1169/49/20230502

the greatest impact on housing prices. The model is tested using the test set, and the intercept term of the model is 0.2901** (3.508), while the goodness of fit is 0.760, slightly higher. The real estate development cost and disposable income are still significant at the 1% significance level. The signs and sizes of the coefficients of the variables are consistent with those of the training set, indicating that the model is reasonable.

Table 3: Regression results of the central-western region model.

	GDP	Real Estate Development Cost	Real estate investment	Disposable income	Rental price
beta (training set)	-0.0747	0.5148**	-0.2251	0.6046**	-0.0567
beta (test set)	-0.2677	0.4515**	-0.1067	0.4737**	-0.4568

4.3. Suggestions and Implications

Based on the above quantitative analysis, the variables that have the greatest impact on housing prices of the eastern region are commercial housing sales area, rental prices, and unemployment rate in descending order. For the central-western Region, the variables are disposable income and real estate development cost in descending order. Based on these influencing factors, the following suggestions are proposed to address the uneven housing prices in different regions and the high housing prices in first-tier cities nationwide.

For the eastern region, first, ensure the growth rate of commercial housing sales area. By regularly analyzing the market operation, ensuring sufficient housing supply, and improving approval services, the growth rate of commercial housing sales area can be ensured to relieve pressure on homebuyers from the supply side. Second, stabilize the rental prices. By stabilizing rental prices, the basic living needs of residents can be ensured. What's more, the rationalization of rental prices will further promote the stability of home prices. The current hot topic of real estate——tax reform，can be used to regulate and influence both the purchase and rental markets [7]. Finally, coordinate the national employment situation. By increasing job opportunities in second and third-tier cities and improving the unemployment insurance system, the pressure on homebuying in first-tier cities can be alleviated.

For the central-western region, first, promote fair and reasonable income distribution. By focusing on the fairness and rationality of income distribution in underdeveloped areas, the attractiveness of talent inflow can be increased, thereby alleviating the imbalance in housing prices in various regions. Second, control real estate development costs. Real estate companies can control development costs through creating economies of scale, system integration, and other means, thus effectively controlling housing prices through transmission mechanisms.

5. Conclusion

In summary, this study first analyzed and determined ten variables that have a significant impact on housing prices, and then conducted empirical analyses on the eastern and central-western regions respectively. The model settings were determined through correlation and multicollinearity tests, and multiple linear regression models were used for regression with the test set used to evaluate the model's performance. Based on the analysis results, the factors that have a significant impact on housing prices in the eastern region are commercial housing sales area, rental prices, and unemployment rate, while in the central-western region, the factors are disposable income and real estate development cost. Based on these conclusions, policy recommendations are proposed. This study makes a certain contribution to exploring the differences in housing prices between the

DOI: 10.54254/2754-1169/49/20230502

eastern and central-western regions of China and the impact of different factors, and provides relevant basis for the government's differentiated macroeconomic policy regulation. It is worth noting that due to the limitations of some data time range and insufficient sample size, the results of this study may not be stable enough. In future research, considerations for time-lag effect can be included, such as using lag terms for variables like real estate development cost, which may make the model setting more reasonable. Additionally, the underlying reasons for regional differences in the factors influencing housing prices also have significant practical implications.

References

[1] *Fang Changchun. Historical and Social Impact of the "98 Housing Reform": Based on the Perspective of Benefit Differentiation. Social Science Digest, 2020, No. 54(06): 59-61.*

[2] *Wang Tongxu. Dynamic Adjustment of First Home Loan Interest Rate Policy. China Business News, 2023-01-10 (002).*

[3] *Zhang Ying, Gui Jin'ai. The Truth and Reflection on the Proportion of China's Land Cost in Housing Price. National Business Situation (Theoretical Research), 2013, No. 1956(06): 24-25.*

[4] *Yi Ying, Liu Meiling. Regional Differences in the Factors Influencing Housing Prices in China. Times Finance, 2019, No. 747(29): 45-46+57.*

[5] *Yu Xiaodong. Is the Current Housing Bubble in Chinese Cities Already Too Large? - An Analysis Based on the Comparison of Renting and Buying Costs in 27 Sample Cities. Economic World, 2021, No. 156(06): 3-9.*

[6] *Xu Huijie. The Influence of Industrial Structure and Population Density on Housing Prices in New Tier-1 Cities. (Master's thesis). Shenyang Jianzhu University, 2022.*

[7] *Zhang Lu. Analysis of Property Tax and Housing Rental Prices. (Master's thesis). Shanghai University of Finance and Economics, 2020.*