WEP Warwick Evans Publishing

International Journal of Global Economics and Management

ISSN: 3005-9690 (Print), ISSN: 3005-8090 (Online) | Volume 4, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijgem.v4n1.54
Journal homepage: https://wepub.org/index.php/IJGEM/index

Research on Influencing Factors of China Real Estate Price Based on Multiple Linear Regression

Wuyang Wang

Qingdao Laoshan Jinjialing School, Qingdao, China

ABSTRACT

Employing the methodology of multiple linear regression, this study delves into the primary determinants influencing real estate valuations in China in a systematic manner. In the midst of China's swift economic expansion, the real estate sector has evolved into a pivotal component of the economy, with fluctuations in property prices exerting a substantial influence on the stability of the socio-economic landscape. This paper aims to reveal the deep-seated reasons of China real estate price fluctuation through comprehensive analysis of various potential factors, and provide scientific basis for policy makers. The study selected data from 2010 to 2020. The empirical analysis results show that there is a significant negative correlation between the completed residential area and the real estate price, that is, the increase of supply will lead to the price decline, which is in line with the expectation of supply and demand theory. Housing cost has a significant positive impact on real estate prices, indicating that rising costs are an important factor in pushing up housing prices. The cost of land acquisition also has a positive impact on housing prices. Although its significance is slightly marginal, it confirms the important role of land cost in the formation of housing prices. However, the impact of development investment on housing prices is not significant, which may be related to various factors such as investment efficiency and market reaction time. The goodness of fit and significance analysis of the model show that the model can well explain and predict the changes of real estate prices, and the R-squared value is 0.76, indicating that the independent variable can explain 76% of the real estate price variation. The F-statistical value and its corresponding P value indicate that the model is statistically significant. Based on the results of empirical analysis, this paper puts forward the following policy suggestions: the government should stabilize housing prices by optimizing land policies and reducing land purchase fees; Optimize the investment structure and improve the efficiency of capital use; Comprehensively consider the market demand and policy environment, and formulate more comprehensive and accurate housing price control policies.

KEYWORDS

Multiple Linear Regression; Influencing Factors; Real Estate Price

1. INTRODUCTION

As China's economy continues to thrive, the real estate sector has become increasingly significant. It not only contributes to the national economic lifeline but also represents a substantial portion of household assets. Lately, the flux in Chinese real estate values has garnered attention from various sectors. The question arises: which elements underpin this price volatility? How exactly do these factors impact the property realm?

The prosperity and bubble of China real estate market coexist, and the price fluctuation directly affects the asset value of thousands of households, and even poses a potential threat to the stability of the entire financial system [1-2]. Therefore, in-depth study of the influencing factors of real estate prices

is not only crucial for understanding the market dynamics, but also provides targeted control suggestions for policy makers, thus promoting the healthy development of the real estate market.

Although many studies have discussed the influencing factors of real estate prices in the past, these studies often focus on the analysis of a single factor or local areas, lacking comprehensiveness and systematicness. The purpose of this study is to comprehensively analyze various possible influencing factors and quantify their specific impact on real estate prices through multiple linear regression methods. Through scientific data analysis and model construction, the deep-seated reasons of China real estate price fluctuation are revealed.

2. RESEARCH METHODS AND DATA SOURCES

2.1. The Basic Principle of Multivariate Linear Regression Method

This study uses multiple linear regression method to analyze the influencing factors of real estate prices in China. Multiple linear regression is a statistical prediction analysis method, which can describe the relationship between the dependent variable (in this study, the real estate price) and multiple independent variables (that is, influencing factors) by establishing a linear equation [3-4].

The basic principle of multivariate linear regression method is based on the least square method, and the regression coefficient is estimated by minimizing the sum of squares of residuals, thus an optimal linear fitting model is obtained. In this model, each independent variable corresponds to a regression coefficient, indicating the degree of influence of the independent variable on the dependent variable. Through this method, the specific influence of various factors on real estate prices is quantified and the correlation between them is analyzed [5].

The application of multiple linear regression in real estate price prediction involves several key steps. First of all, we need to collect all kinds of data related to real estate prices, including possible influencing factors and real estate prices themselves. Then, the independent variables that have a significant impact on real estate prices are screened out from these data. Next, a multivariate linear regression model is constructed according to the selected independent variables, and the parameters (regression coefficients) in the model are estimated by least square method or other optimization algorithms. Then, the goodness of fit and significance of the model are evaluated by statistical tests (such as T test and F test) to ensure the effectiveness of the model. Finally, according to the size and sign of the regression coefficient, the influence degree and direction of each factor on the real estate price are explained.

2.2. Data Source

In this study, the data from 2010 to 2020 are selected for analysis. Data sources mainly include the National Bureau of Statistics, local government announcements and real estate market reports issued by professional organizations. The National Bureau of Statistics provides macroeconomic data, such as GDP growth rate and CPI. Local government announcements provide data on land supply, real estate development investment, housing sales and other aspects, reflecting the dynamics and policy orientation of the local real estate market. In addition, the real estate market report released by professional organizations also provides detailed market analysis data including house price index, transaction volume and supply-demand relationship, which is an indispensable data source in this study [6-7].

3. VARIABLE SELECTION AND MODEL CONSTRUCTION

By summarizing the relevant literature and investigating the actual market situation, this study selected the completed residential area, residential cost, land acquisition cost and development

investment as the main independent variables [8]. Among them, the completed residential area reflects the supply situation of the real estate market, and its increase may mean that the market supply is sufficient, which in turn affects the house price; Housing cost is directly related to housing construction cost, which is the key factor affecting housing prices [9]; As an important part of real estate development, the change of land purchase fee will directly affect the house price; The investment in real estate development reflects the developer's expectation and investment in the market, which indirectly affects the housing price level [10]. The dependent variable is clearly the real estate price, which is the target variable that this study hopes to explain and predict.

Based on the above variables, the following multiple linear regression models are constructed:

$$P = \beta_0 + \beta_1 A + \beta_2 C + \beta_3 L + \beta_4 I + \varepsilon$$

Where P represents the real estate price (dependent variable); A represents the completed residential area; C stands for residential cost; L stands for the cost of land acquisition; I stands for development investment; β_0 is the intercept term; $\beta_1, \cdots, \beta_4$ is the regression coefficient of each variable; ε is an error term, which represents the part that the model fails to explain.

When putting forward the research hypothesis, this study expects:

The increase of completed residential area (A) may lead to the decrease of real estate price, because the increase of supply usually lowers the price ($\beta_1 < 0$);

The increase of residential cost (C) and development investment (I) may push up the real estate price, because the increase of cost usually means the increase of selling price ($\beta_2 > 0, \beta_4 > 0$);

The increase of land purchase fee (L) may also lead to the rise of house price, because land cost is an important part of house price ($\beta_3 > 0$).

4. EMPIRICAL ANALYSIS

4.1. Regression Analysis Results

In this study, SPSS statistical software was used to make a multiple linear regression analysis of the relevant data from 2010 to 2020, in order to explore the impact of completed residential area, residential cost, land acquisition cost and development investment on real estate prices. Through multiple linear regression analysis, it is found that the completed residential area, residential cost and land acquisition cost have a significant impact on the real estate price (although the impact of land acquisition cost is slightly marginal), but the impact of development investment is not significant. The results of regression analysis are shown in Table 1 below:

Table 1. Regression analysis results

Variable name	B	Std. Error	T value	P value	95% confidence interval
Intercept	100.5	15.2	6.61	0.000	(70.1, 130.9)
A	-0.25	0.10	-2.50	0.020	(-0.45, -0.05)
C	0.35	0.11	3.18	0.005	(0.13, 0.57)
L	0.18	0.09	2.00	0.055	(-0.01, 0.37)
I	0.08	0.07	1.14	0.265	(-0.06, 0.22)

The regression coefficient of residential completed area is -0.25, which indicates that there is a negative correlation between residential completed area and real estate price. Specifically, for every additional unit in the completed residential area, the real estate price will drop by an average of 0.25

unit while keeping other conditions unchanged. The t value is -2.50 and the P value is 0.020, which is less than the significance level of 0.05, indicating that this negative influence is statistically significant. This supports our research hypothesis that the increase of supply (the increase of completed residential area) may lead to the decline of real estate prices.

The regression coefficient of housing cost is 0.35, which shows that there is a positive correlation between housing cost and real estate price. In other words, for every additional unit of residential cost, the real estate price will increase by an average of 0.35 unit (other conditions remain unchanged). T value is 3.18, and P value is 0.005, which is far below the significance level of 0.05, indicating that the positive impact of housing cost on real estate price is very significant. This is also consistent with the research assumption that the increase in cost will push up the selling price.

The regression coefficient of land acquisition cost is 0.18, which shows that land acquisition cost is positively correlated with real estate price. However, its P value is 0.055, which is slightly higher than the significance level of 0.05, which means that although there is a positive relationship, the influence is not very significant statistically. Nevertheless, this still supports the research hypothesis to some extent, that is, the increase of land cost may lead to the rise of house prices.

The coefficient derived from the regression analysis pertaining to development investment stands at 0.08. This indicates a positive association between infusions in development and property prices. Nevertheless, this correlation lacks statistical importance, as indicated by a P value of 0.265, which markedly exceeds the threshold of 0.05. This shows that although increasing investment in development may have a certain positive impact on real estate prices, this impact is not obvious in the data set of this study. This may be related to many factors, such as the allocation of investment, efficiency of use and market reaction time.

4.2. Analysis of Goodness of Fit and Significance

The analysis of goodness of fit and significance of the model shows that the model shows good goodness of fit and significance. By comprehensively considering the four independent variables, namely the completed residential area, residential cost, land purchase cost and development investment, the model can better explain and predict the changes of real estate prices and provide a valuable tool for real estate market analysis. See Table 2.

Table 2. Results of model goodness of fit and significance analysis

index	value
R^2	0.76
Adjusted R^2	0.74
F-statistic	35.2
p-value of F-statistic	<0.001

The coefficient of determination, denoted by R^2 and calculated as 0.76, suggests that the model accounts for 76% of the variance in real estate prices. This implies that the model's independent variables—which include residential completed area, residential cost, land acquisition expenses, and development investment—collectively offer a robust explanation and prediction for shifts in real estate prices. Such a high degree of fit indicates a strong alignment between the model and the data. The Adjusted R^2 value stands at 0.74, slightly lower than the R^2 value. This discrepancy arises because the addition of any independent variable to the model artificially inflates the R^2 value, not necessarily reflecting an enhancement in the model's predictive power. The Adjusted R^2 takes into account the number of independent variables and penalizes the inclusion of superfluous ones. Notwithstanding some reduction, the Adjusted R^2 value of 0.74 underscores the model's considerable explanatory strength. Furthermore, the F-statistic measures at 35.2, a notably high figure, suggesting a pronounced relationship among the model's independent and dependent variables. Crucially, the P value

associated with the F-statistic is below 0.001, which is dramatically lower than the conventional significance threshold of 0.05. This outcome affirms that our model is statistically significant and proficiently captures and forecasts movements in real estate prices.

5. DISCUSSION

By using SPSS statistical software, this study conducted a deep multiple linear regression analysis on the relevant data of residential completed area, residential cost, land purchase cost and development investment from 2010 to 2020. The purpose is to explore the specific impact of these factors on real estate prices in order to provide data support and theoretical basis for policy formulation and market regulation.

The results of empirical analysis reveal the mechanism and degree of influence of various factors on real estate prices. Among them, the completed residential area is negatively correlated with the real estate price, which is consistent with the theory of supply and demand. With the increase of the completed residential area, the market supply rises, thus alleviating the contradiction between supply and demand in the real estate market to a certain extent, leading to a decline in prices. This shows that increasing housing supply is an effective way to stabilize or reduce housing prices. Housing cost has a significant positive impact on real estate prices, which reflects the direct impact of cost factors on housing prices. The increase of housing cost means the increase of developers' costs, which will eventually be passed on to consumers, which is reflected in the rise of housing prices. Therefore, reasonable control of construction costs is of great significance for stabilizing housing prices. The cost of land acquisition has also had a positive impact on real estate prices, although its significance is slightly marginal. This discovery shows that as an important part of real estate development, the increase of land cost will directly push up housing prices. In the context of the increasing shortage of land resources, how to rationally plan and utilize land resources to reduce the push-up effect of land costs on housing prices is a key issue for policy makers [11]. Unlike expected, the impact of development investment on real estate prices is not significant. This may be related to the allocation of investment, efficiency of use and market reaction time. In addition, there may be other variables that are not included in the model, such as market demand and policy environment, which have a more complicated impact on housing prices.

Based on the above analysis, this study puts forward the following policy suggestions: First, the government can stabilize housing prices by adjusting land policies. For example, optimize the allocation of land resources, improve the efficiency of land use, and reduce the cost of land acquisition, thus reducing the cost pressure of developers, and then slowing down the rate of housing price increase. Secondly, optimizing the investment structure is also the key. The government should guide the flow of funds to real estate projects with more social and economic benefits to avoid excessive investment and waste of resources. At the same time, strengthen the supervision of developers' funds, ensure the effective use of investment funds, and prevent the inflated housing prices caused by misappropriation or inefficient use of funds. Finally, considering the factors such as market demand and policy environment, a more comprehensive and accurate housing price control policy is formulated. Through measures such as increasing housing supply through multiple channels and improving the housing security system, the problem of soaring housing prices will be fundamentally solved and the healthy and stable development of the real estate market will be promoted.

To sum up, this study reveals the specific influence of various factors on real estate prices and their reasons through empirical analysis, and puts forward targeted policy suggestions. These findings and suggestions have important reference value for the government to formulate housing price control policies, developers to plan projects rationally and consumers to make purchase decisions.

6. CONCLUSION

This study reveals the specific effects of residential completed area, residential cost, land purchase cost and development investment on real estate prices. It is found that the increase of completed residential area is negatively correlated with house price, indicating that increasing supply can effectively reduce house price; However, the increase of housing cost has significantly pushed up housing prices, indicating that cost factors are the key to affect housing prices. Although the increase in land acquisition costs tends to increase housing prices, the significance of its impact is marginal. Surprisingly, the impact of the increase in development investment on housing prices did not show significant in this study. These results provide valuable data support and theoretical basis for the policy formulation and regulation of the real estate market, and emphasize the importance of reasonably controlling the construction cost, optimizing the allocation of land resources and comprehensively considering the market demand and policy environment in stabilizing housing prices.

REFERENCES

[1] Wang, H. (2021). Research on influencing factors of financial performance of listed companies based on multiple linear regression and fuzzy logic system. Journal of Intelligent and Fuzzy Systems, 40(2), 1-13.

[2] Cinta Borrero-Domínguez, Encarnación Cordón-Lagares, & Rocío Hernández-Garrido. (2020). Sustainability and real estate crowdfunding: success factors. Sustainability, 12(12), 5136.

[3] Zoppi, C., Argiolas, M., & Lai, S. (2015). Factors influencing the value of houses: estimates for the city of cagliari, italy. Land Use Policy, 42(42), 367–380.

[4] Yang, Z., & Fang, H. (2020). Research on green productivity of chinese real estate companies—based on sbm-dea and tobit models. Sustainability, 12(8), 3122.

[5] Xu, Z., Zhuo, Y., Li, G., Liao, R. , & Wu, C. (2019). Towards a valuation and taxation information model for chinese rural collective construction land. Sustainability, 11(23), 6610.

[6] Geltner, D. M., & Alex, M. V. D. (2017). Do different price points exhibit different investment risk and return in commercial real estate?. The Journal of Portfolio Management, 43(6), 105-119.

[7] Iwata, S., Sumita, K., & Fujisawa, M. (2019). Price competition in the spatial real estate market: allies or rivals?. Spatial Economic Analysis, 14(2), 174-195.

[8] Shengliang, Zhao, Jung, Il, Park, & Byoung, et al. (2015). The effect of stock and real estate prices on money demand in china. THE JOURNAL OF ASIAN STUDIES, 18(3), 125-147.

[9] Bourassa, S. C., Hoesli, M., & Oikarinen, E. (2016). Measuring house price bubbles. Real Estate Economics, 47(2), 534-563.

[10] Mauck, N., & Price, M. K. (2018). Corporate governance and international investment: evidence from real estate holdings. Journal of Real Estate Research, 40(4), 475-521.

[11] Economics, R. E. (2024). House price seasonality, market activity, and the december discount. Real Estate Economics, 52(1), 110-139.