

# Analysis of Z City House Price Forecast based on Application of Multiple Linear Regression

Kun Zhang<sup>a</sup>, Xin Zhu<sup>b</sup>

School of Economics and Management, Guangxi University of Science and Technology, Liuzhou 545006, China

<sup>a</sup>1169360463@qq.com, <sup>b</sup>5266736@qq.com

**Keywords:** House price, national economy, Multiple linear regression.

**Abstract:** House prices are a barometer of the real estate industry and a barometer of the national economy. The change in housing prices is not only the focus of the real estate industry, but also the focus of the state and the government as well as the public. For the Chinese who pay attention to starting a family and re-establishing their business, the house is the foundation of the world. Therefore, the Chinese people are more concerned about the fluctuations in housing prices. This paper uses multiple linear regression analysis to select four indicators that have a significant impact on Z city house prices from the nine indicators that have an impact on housing price changes. By establishing a new multiple regression model, the short-term fluctuations of Z city house prices are predicted.

## 1. Introduction

With the development of the market economy, the real estate industry is also constantly improving. The real estate industry with the residential real estate industry as the core has become more and more important in the national economy and has gradually become the pillar industry of China's economic development. Therefore, the real estate industry has received the attention of many economists. More and more economists are committed to exploring the unique laws of the development of China's real estate industry, trying to find the recognized factors that affect housing price changes and establishing short-term fluctuations in housing prices.

Well-known scholars believe that from the long-term trend of housing prices, the development of China's western region relative to the financing market in the Middle East is not perfect, relatively single, the country's credit policy has a greater impact on housing prices in the region, and the central and eastern regions Changes in housing prices are more dependent on changes in economic growth and demand in the region[1].Baishuang (2008) proposed that the main influencing factors of real estate prices are consumer purchasing power, product cost, number of enterprises, per capita resource possession, investment status and tax policy[2].

In 2009, the Z city staged a recovery, recovery, and rising situation. From the downturn in the property market in January to the full recovery of real estate in May, and the rapid rise in housing prices after September and the frequent occurrence of "land kings" in December, the property market showed a trend of "increasing prices". According to the statistics of the property market of the Z City Housing Authority, in 2009, the average selling price of commercial housing in Z City was 4,550 yuan/m<sup>2</sup>. The sales area of commercial housing in the city was 11.989 million square meters, an increase of 62.7% over the previous year. Among them, the residential sales area was 10.851 million square meters, an increase of 64.8% over the previous year; the sales of commercial housing were 51.48 billion yuan, an increase of 77.9% over the previous year. Sales amounted to 43.99 billion yuan, an increase of 85.7% over the previous year. In terms of regional price, in 2009, the sales of 4,000 yuan / square meter - 5,000 yuan / square meter of commercial housing and 5,000 yuan / square meter ~ 6000 yuan / square meter of commercial housing accounted for 2 / 3 of the total number of commercial housing sales It is the main transaction price of Z City, and its proportion is greatly improved compared with the previous two years, so that the momentum of Z home price rise

can be fully felt. As far as buyers are concerned, the proportion of buyers from the “other regions of the province” in Z City in 2009 exceeded half of the total number of commercial housing sales in Z, reaching 52%, and the number of purchases reached 58285, an increase of 118.9%. At the same time, the proportion of buyers from other provinces and cities and suburbs of this city has also increased. This aspect shows that with the construction and development of Z City, there are more and more foreigners working and living in Zhengzhou, and they have more urgent needs for housing. On the other hand, as a regional central city, Z City has become a suburb. The preferred target for wealthy people in counties and provinces, and Z City as a second- and third-tier city has also become an investment target for some “speculators”. The average price of local buyers in Z city is 4,288 yuan / square meter, which is the lowest purchase price among people in different regions; the average purchase price of buyers in other regions exceeds 4,700 yuan / square meter. This is mainly because buyers in the “city of the city” can purchase affordable housing at a lower price. However, the increase in the purchase price of “the city's urban area” is the highest. In 2009, the real estate market was “hot”, especially in the situation that there was a certain degree of “panic buying” at the end of the year, and some “city of the city” buyers no longer Waiting for affordable housing, and joining the purchase of commercial housing.

In China's economic development, policy factors have a great impact on it. It is precisely because of this that China's real estate market development shows an obvious stage. If you use a large span of data, you will lose a lot of short-term information, so this article comprehensive Considering the representativeness of the variables and the availability and reliability of the data, the following indicators have been selected for the implementation of the housing system reform, that is, the data from year to month, and the real estate prices are quantitatively studied to reveal its internal operating mechanism[3]. According to the use, real estate includes commercial real estate, residential real estate, office real estate and other categories. The prices of several types of real estate vary greatly, and it is impossible to put these types of real estate together in general terms. Since the price of residential real estate is of the utmost concern to the public, the article conducts the next empirical study based on the real estate price of the home. Since the statistical bureaus have not released the housing sales price index of Z city in recent years, the average transaction price of the house, that is, the total residential sales value divided by the residential sales area, is used to measure the house price fluctuation in Z city.

## **2. The Basic Principle of Multiple Linear Regression Model**

The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph). All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use.

In mathematical modeling, the most basic one is the linear regression model, and there is only one influencing factor, namely the independent variable. However, in real life, the situation is more complicated. The occurrence of a phenomenon cannot be only one or only one. The commonly used ones in our statistics are multiple linear regression models and linear regression with two or more independent variables. model.

The general representation is as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad i=1,2,\dots,n \quad (1)$$

Where Y is the dependent variable,  $X_1, X_2, \dots, X_k$  are independent variables,  $\beta_0, \beta_1, \beta_2, \dots, \beta_k$  are the parameters to be estimated, that is, the regression coefficient,  $\mu$  is the random error term, and k is the number of independent variables. i is the subscript of the observation and n is the sample size.

Basic assumptions:

1 The independent variable is random or fixed (no multicollinearity).

2 Random error term with characteristics of mean 0, same variance, and no sequence correlation.

That is,  $E(\mu_i)=0$ ,  $\text{Var}(\mu_i)=E(\mu_i^2)=\sigma^2$ ,  $\text{Cov}(\mu_i, \mu_j)=E(\mu_i, \mu_j)=0, i \neq j, 1, 2, \dots, n$

3 dependent variables are not related to randomness

That is,  $\text{Cov}(X_{ji}, \mu_i)=0, j=1, 2, \dots, k$

4 random items satisfy the normal distribution

That is,  $\mu_i \sim N(0, \sigma^2)$ .

### 3. Indicator Selection and Data Source

In order to understand the correlation between housing prices and commodity housing factors, this paper uses a multiple linear regression model, combined with Z city price data from 2003 to 2016 for quantitative analysis. First, this paper selects nine representative indicators that have an important impact on housing prices: average price of commercial housing (Y), regional ( $X_1$ ), per capita wage ( $X_2$ ), household registration population urbanization ( $X_3$ ), population density ( $X_4$ ), real estate investment ( $X_5$ ), completed area ( $X_6$ ), construction and installation costs ( $X_7$ ), and medium and long-term loan balance ( $X_8$ ). This paper selects relevant data for 2003-2016, the data is shown in Table 1.

Table 1. Factors affecting the average price of residential commercial housing sales in Z City

YEAR	Y	X1	X2	X3	X4	X5	X6	X7	X8
2003	2318	1102.28	13537.99	21.93	887.82	74.2644	277.62	112	213.27
2004	2639	1377.89	15023.84	24.53	901.36	121.1829	279.36	113.1	238
2005	2966	1660.6	16695.7	25.62	912.84	168.1013	362.96	126.87	277
2006	3232	2013.48	18862.08	26.34	928.85	229.8933	411.81	144.3	368
2007	3552	2486.75	22156.23	26.97	949.52	298.7582	538.57	144.43	480
2008	3994	3003.99	26477.63	26.96	966.44	429.9506	626.97	194.99	528
2009	4580	3308.51	29836.72	27.33	982.37	513.8302	530.19	234.69	733
2010	5736	4040.89	32777.81	30.68	1293.31	775.1607	751.66	273.54	935
2011	7562	4912.7	35755.56	31.79	1356.57	923.6	1304.76	308.14	1073
2012	7687	5547	40471.52	32	1440.37	1095.137	1043.2	330.45	1197
2013	9266	6201.85	45066	32.5	1234.35	1445.328	760.56	369.14	1461
2014	9519	6776.99	49756.03	47.27	1259.47	1743.513	1122.87	421.32	1737
2015	10010	7315.19	52987	48.25	1087.83	2000.195	670.5	438.31	2100
2016	23423	7994.16	56856.16	49.32	1025.45	2290.8	1056.28	467.3	2668

### 4. Model Analysis and Inspection

By analysis, we can get that the P values of the four variables T test of  $X_3, X_4, X_6$  and  $X_7$  are all greater than 0.05, which means that the coefficients of these four variables are not significantly affected by the model and are excluded. Therefore, we can get the following multiple linear regression models:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_5 X_5 + \beta_8 X_8 + \mu_i \quad (2)$$

According to the analysis, GDP, AS, EIA) and medium and long-term loans the standard regression coefficients corresponding to the four variables of LOA are 3.034, -0.313, -4.827 and 5.966, respectively. The final regression model is:

$$Y=3.034X_1-0.313X_2-4.827X_5+5.966X_8+2069.981 \quad (3)$$

Among them, Y represents the average price of residential housing in Z city,  $X_1$  represents regional GDP,  $X_2$  represents per capita wage,  $X_5$  represents real estate investment, and  $X_8$  represents medium and long-term loan balance.

According to the multiple linear regression equation we obtained above, the average selling price of commercial housing in Z city from 2003 to 2016 is predicted, and the predicted value is compared with the actual value. It is found that the Z city commercial residential house is estimated by using the multiple linear regression equation obtained. The average relative error between the actual value and the predicted value of the average selling price is 5.57%, and the accuracy is a high level, which can be used to predict the average selling price of commercial housing in the future years of the region.

## 5. Summary

In this paper, the method of multiple linear regression analysis is adopted, and eight independent variables that have significant influence on commercial housing are selected. Through the gradual regression, four variables that have a significant impact on housing prices are selected, and the commercial housing of Z city is established. The optimal linear regression model of the change in average selling price is validated to some extent by inspection and analysis. At the same time, there are certain deficiencies in the research of this paper. The indicators selected in this paper are quantifiable data that can be obtained, so that the variables that may be important are ignored to some extent, which affects the scientific and rigor of the research. Moreover, the data in this paper comes from different statistical websites, which will make a difference between the statistical data and the actual value, which expands the difference between the actual value and the predicted value of the house price to some extent.

## References

- [1]. LIANG Yunfang, GAO Tiemei. An Empirical Analysis of Regional Differences in China's Real Estate Price Volatility[J]. Economic Research, 2007, (08): 133-142.
- [2]. Baishuang. Analysis of the Determinants of Real Estate Price——An Empirical Study of Panel Data in 31 Regions of China[J]. Studies in Financial and Economic Issues, 2008, (08): 107-110.
- [3]. Wang Wei. Application of Logistics Regression Model in Forecasting Housing Price in Guangdong Province[J]. Modern Business and Industry,2010,22(16):304-306.