Analysis of Influential Factors of Housing Price in Nanjing
Based on Improved Grey Relational Analysis

Qingguo Tang* and Peng Jin
School of Economics and Management, Nanjing University of Science and Technology,
China, 210094

*Corresponding author: tangqguo@163.com

Keywords: Housing price, Influential factors, Grey correlation.

Abstract. This paper based on the real estate related data of Nanjing City from 2006 to 2015, selects the index of influencing factors of house price appropriately and introduces the method of grey absolute relational analysis to construct the grey relational degree model of house price influencing factors, and quantitatively calculate the grey correlation between each index and house price. The results show that the two factors of urban GDP and residents’ annual income have the greatest impact on house prices, while the investment of commercial buildings, the sales area of commercial buildings and the completed area of commercial houses have less impact.

Introduction

Since the readjustment of the land-related laws and regulations in 1978, China's real estate industry has gone through 30 years of stormy history along with the reform and opening-up. In particular, China’s real estate industry experienced a rapid boom in the past decade, the first round of upswings in China's real estate market started in the early 2000s. Real estate industry as a basic and leading industry. Since the real estate products provided by the real estate industry are both living materials and means of production, they are indispensable basic elements of social and economic activities. All sectors of the national economy can not do without real estate. Therefore, it can be said that the real estate industry occupies an important position in the national economy as well as in the personal economy. National Bureau of Statistics data show that in 2016 the real estate industry directly accounted for 6.5% of GDP, driven by GDP growth of about 0.22 percentage points, the real estate industry chain together accounted for about 13% of GDP. Data from the China Household Financial Survey (CHFS) shows that housing is the most important component of household wealth. Among per capita wealth of families in the country, the proportion of real estate is 65.99%.

The grey relational analysis method is an active multi-factor statistical analysis method, the system theory has been a lot of improvements after years of research and development, no matter from the improvement of theoretical thought or the application of this analytical method have made many achievements. In terms of grey relational theory, in the theory of grey relational degree, Deng[1-3] proposed the grey system theory, and gave a general method of calculating the degree of correlation. Mei[4] proposed the concept and calculation method of absolute grey relation degree in order to overcome the deficiency of general relation degree. Wang[5] fully considered the positive and negative correlations between things and factors, and gave a method of calculating the correlation coefficient and the degree of correlation. In order to overcome Deng's correlation and the lack of absolute relevance, Tang[6] proposed T-relatedness. In the continuous development of grey relational analysis, Chen and Wei.[7] analyzed the defects of related degrees in the past and proposed a new method of calculating the degree of correlation.

For Chinese society, the volatility of housing prices will cause a series of economic problems. Therefore, many scholars have conducted extensive and in-depth studies on the factors affecting housing prices. Wen[8] developed a simultaneous-equations model to explore the interaction between housing price and land price. This model uses urban land price and housing price as endogenous variables and five factors for land price and seven factors for housing price as exogenous variables,
and as a whole, housing price has greater influence on land price. Zhang et al.
[9] apply a non-linear modeling approach, the Nonlinear Auto Regressive Moving Average with eXogenous inputs (NARMAX), to investigate determinants of housing prices in China. Estimation results mainly identify some key monetary and price variables in interpreting housing price dynamics, including most notably mortgage rate, producer price, broad money supply and real effective exchange rate.

The Factors Influencing Housing Price in Nanjing City

Variable Selection

Housing prices are mainly related to supply and demand factors. We choose the following influencing factors: urban GDP, urban resident population, annual income of urban residents, urbanization rate, investment in residential houses, sales of residential houses, completion of commercial residential houses, Commercial housing sales area. We use the above eight factors for grey relational analysis, the eight influencing factors are represented by \( X_1, X_2, \ldots, X_8 \) respectively, \( X_9 \) represents the housing price.

Data Selection and Source

This paper selects the data of ten years from 2006 to 2015. The data mainly comes from the statistical bulletin of urban residents and social development in Nanjing, the Nanjing Real Estate Market Annual Report and the website of Nanjing Bureau of Statistics. The results of the data are summarized in Table 1.

Table 1. The data of ten influencing factors from 2006 to 2015.

| year | \( X_0 \) | \( X_1 \) | \( X_2 \) | \( X_3 \) | \( X_4 \) | \( X_5 \) | \( X_6 \) | \( X_7 \) | \( X_8 \) |
|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 2006 | 6935      | 2774      | 719       | 17538     | 0.764     | 3120.1    | 632.6     | 351.2     | 1246.5    |
| 2007 | 7632      | 3284      | 741       | 20317     | 0.768     | 3582.7    | 754.2     | 445.9     | 1574.9    |
| 2008 | 8310      | 3815      | 759       | 23123     | 0.770     | 4097.7    | 743.3     | 508.2     | 1598.7    |
| 2009 | 9607      | 4230      | 771       | 25504     | 0.772     | 4570.4    | 872.5     | 595.7     | 1876.2    |
| 2010 | 13619     | 5013      | 801       | 28312     | 0.770     | 5664.2    | 943.2     | 754.8     | 1976.0    |
| 2011 | 14981     | 6146      | 811       | 32200     | 0.785     | 4366.0    | 808.8     | 896.7     | 1243.5    |
| 2012 | 14460     | 7202      | 816       | 36322     | 0.797     | 6050.0    | 1022.8    | 1105.7    | 2074.6    |
| 2013 | 16435     | 8012      | 819       | 39881     | 0.805     | 6572.9    | 1222.0    | 1120.1    | 2243.4    |
| 2014 | 18004     | 8821      | 822       | 42568     | 0.809     | 6892.3    | 1345.2    | 1176.2    | 2419.2    |
| 2015 | 18697     | 9721      | 824       | 46104     | 0.813     | 7565.5    | 1542.5    | 1235.4    | 2766.4    |

Note: In the above table, the unit of \( X_0 \) and \( X_5 \) is yuan, the unit of \( X_1, X_2, \ldots, X_8 \) is billions of yuan, the unit of \( X_2 \) is ten thousand, and the unit of \( X_3, X_4, X_6, X_7 \) and \( X_8 \) is ten thousand square meters.

Table 2. The standardized data for Table 1.

| year | \( X_0 \) | \( X_1 \) | \( X_2 \) | \( X_3 \) | \( X_4 \) | \( X_5 \) | \( X_6 \) | \( X_7 \) | \( X_8 \) |
|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 2006 | 1.000     | 1.000     | 1.000     | 1.000     | 1.000     | 1.039     | 1.000     | 1.002     |
| 2007 | 1.101     | 1.184     | 1.158     | 1.105     | 1.148     | 1.239     | 1.270     | 1.267     |
| 2008 | 1.198     | 1.375     | 1.056     | 1.318     | 1.008     | 1.313     | 1.221     | 1.447     | 1.286     |
| 2009 | 1.385     | 1.525     | 1.072     | 1.454     | 1.010     | 1.399     | 1.433     | 1.696     | 1.509     |
| 2010 | 1.964     | 1.807     | 1.114     | 1.614     | 1.027     | 1.465     | 1.549     | 2.149     | 1.589     |
| 2011 | 2.160     | 2.216     | 1.128     | 1.836     | 1.043     | 1.815     | 1.000     | 2.553     | 1.000     |
| 2012 | 2.085     | 2.596     | 1.135     | 2.071     | 1.050     | 1.939     | 1.680     | 2.892     | 1.668     |
| 2013 | 2.370     | 2.888     | 1.139     | 2.274     | 1.054     | 2.107     | 2.007     | 3.189     | 1.804     |
| 2014 | 2.596     | 3.180     | 1.143     | 2.427     | 1.059     | 2.209     | 2.210     | 3.349     | 1.945     |
| 2015 | 2.696     | 3.504     | 1.146     | 2.629     | 1.064     | 2.425     | 2.534     | 3.518     | 2.225     |
Dimensionless Data Processing

Since the eight indicators which we are selected contain various types of data, the meaning of each factor and the dimension of the data are not the same. Therefore, before data analysis, we usually need to standardize the data and use the standardized data for data analysis. The standardization of data mainly includes two aspects of data processing and dimensionless treatment. This paper uses the dimensionless method to deals with the data in table 1. The results are shown in table 2.

Improved Grey Relational Model

Grey Relational Model

Grey relational analysis is mainly for the quantitative analysis of the data in the dynamic development of the system. It is based on the degree of similarity or dissimilarity between the development trends of the factors to measure the degree of the closeness between the factors, which is essentially the degree of each evaluation object is close to the ideal object. The closer the evaluation object is to the ideal object, the greater the degree of its association. Among them, the evaluation object with the highest degree of correlation is the best. Therefore, the order of correlation can be used to sort the objects to be evaluated and compared. The steps of using grey relational degree to evaluate the comprehensive evaluation are as follows:

1) Tabulate the indicators of all the evaluated objects.
2) Since there is no computing relationship between the data of each index, therefore, the data must be dimensionless.
3) Construct the ideal object, that is, the optimal value of each index in the non-dimensional evaluation object as the ideal target value.
4) Calculate the index correlation coefficient, which is calculated as follows:

\[
\xi_i(k) = \frac{\Delta_{\text{min}} + \rho \Delta_{\text{max}}}{\Delta_{i}(k) + \rho \Delta_{\text{max}}}, \quad \Delta_{\text{min}} = \min_{i,k} [x_i(k) - x_i(k)], \quad \Delta_{\text{max}} = \max_{i,k} [x_i(k) - x_i(k)],
\]

\[
\Delta_i(k) = [x_i(k) - x_i(k)], \quad i = 1, \cdots, n; k = 1, \cdots, m
\]

Where, \(n\) is the number of evaluation objects, \(m\) is the number of evaluation object indicators, \(\xi_i(k)\) is the correlation coefficient between the kth index of the ith object at the same index of the ideal object; \(\Delta_{\text{min}}\) represents the minimum absolute difference between the kth index value of each evaluation object and the kth index value of the ideal object, and then according to \(i = 1, \cdots, n\) find the smallest of all the smallest absolute difference; \(\Delta_{\text{max}}\) represents the maximum absolute difference between the kth index value of the evaluation object and the kth index value of the ideal object, then according to \(i = 1, \cdots, n\) find the maximum of all the largest absolute difference; \(\Delta_i(k)\) is the absolute difference between the kth index value of the evaluation object and the kth index value of the ideal object. \(\rho\) is the resolution coefficient, the smaller the \(\rho\) is, the larger the resolution is. Generally, the value interval of \(\rho\) is \([0, 1]\). In order to simplify the calculation, \(\rho = 0.5\) is generally taken.

5) Establish the AHP model.
6) Determine the judgment matrix, calculate the weighting coefficients and the degree of weighted correlation at each level, and the formula of weighted degree of correlation is:

\[
\gamma_i = \sum_{k}^{m} \xi_i(k) \omega_k
\]

Where \(\gamma_i\) is the weighted relevance of the ith evaluation object to the ideal object and \(\omega_k\) is the weight of the kth indicator.

7) According to the degree of weighted relevance, each evaluation object is sorted, establishing the correlation order of evaluation object, so that the object with larger correlation degree can be drawn. The bigger the correlation degree is, the better the comprehensive evaluation result will be.
It can be seen from the above steps of calculating the degree of correlation that the degree of relevance has the following two defects:

(1) Affected by the minimum and maximum absolute differences, the value of the degree of association will be affected once an extreme maximum or minimum point appears in the selected data.

(2) The correlation coefficient and the correlation degree are affected by the resolution coefficient. In the actual calculation, the value of $\rho$ is generally 0.5, which is not universal and general.

**Improved Grey Correlation Model**

Let $X_0, X_1, \cdots, X_m$ be $m+1$ factors, the observed values of each factor are:

$X_0 = (x_0(1), x_0(2), \cdots, x_0(n))$, $X_1 = (x_1(1), x_1(2), \cdots, x_1(n))$, $\cdots$, $X_m = (x_m(1), x_m(2), \cdots, x_m(n))$.

(1) Find the initial point initialization

\[ X^0_0 = (x_0(1) - x_0(1), x_0(2) - x_0(1), \cdots, x_0(n) - x_0(1)) = (x^0_0(1), x^0_0(2), \cdots, x^0_0(n)), \]

\[ X^0_i = (x_i(1) - x_i(1), x_i(2) - x_i(1), \cdots, x_i(n) - x_i(1)) = (x^0_i(1), x^0_i(2), \cdots, x^0_i(n)), \cdots, \]

Where $X^0_0 = X^1_0 = 0$.

(2) Solve \[ |S_0|, |S_i|, |S_i - S_0|, |S_0| = \left| \sum_{k=1}^{n-1} x^0_k(k) + \frac{1}{2} x^0_n(n) \right|, \]

\[ |S_i - S_0| = \left| \sum_{k=2}^{n-1} (x^0_i(k) - x^0_k(k)) + \frac{1}{2} (x^0_i(n) - x^0_k(n)) \right|. \]

(3) The grey absolute relation of $X_0$ related to $X_i$ is:

\[ \varepsilon_{ai} = \frac{1 + |S_0| + |S_i|}{1 + |S_0| + |S_i| + |S_i - S_0|}. \]

According to the above steps (1), (2) and (3), similarly, we can get the absolute degree of $X_0$ relative to $X_2, X_3, \cdots, X_m$.

**Specific Calculations and Results Analysis**

Through Matlab software, we can get the absolute degree of $X_0$ relative to $X_1, X_2, \cdots, X_m$. The calculation results of absolute relevance are shown in Table 3

| Factors | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|
| Correlation | 0.9235 | 0.8343 | 0.8451 | 0.7929 | 0.6828 | 0.7226 | 0.7248 | 0.7324 |
| Rank | 1     | 3     | 2     | 4     | 8     | 7     | 6     | 5     |

From the above table, we can see the value of grey absolute relevance and the ranking, in this paper, the factors that affect the housing price, the GDP rankings the most front. GDP is one of the indicators that reflects the level of macroeconomic development and the real estate industry to develop economic background. When the economic situation is optimistic, it will attract more investment, it will also inspire people to increase the purchase of housing. People usually think that the economic growth situation is better, the greater the possibility of rising house prices. While the economic prosperity will also encourage developers to expand the scale of investment and increase the supply of housing, so GDP and demand from the supply of the housing market are rising.

The second factor is the annual income of residents. The per capita disposable income of urban residents can more directly reflect the purchasing power and consumption level of consumers in a region than the total GDP of a region. With the increase of residents' income, the standard of living
has been greatly improved, willingness to spend on housing and consumption power will be increased accordingly. In general, other conditions remain unchanged, the residents income and residential demand will be the same change.

Other influencing factors from descending order are: urban population, urbanization rate, sales of commercial buildings, investment in commercial buildings, sales area of commercial buildings and completed area of commercial buildings.

Conclusions

This paper selects Nanjing real estate related data from 2006 to 2015, we choose eight economic factors that affect the housing price and use the improved grey correlation analysis to analyze the above factors. The results show that the annual income of urban GDP and residents have the greatest impact on the housing prices; while the sales area of commercial housing and the completion of commercial housing have the weakest impact on housing prices.

References

[1] Y. Deng. Basic method for grey system (first edition). 1987.
[2] Y. Deng. Grey prediction and decision. China University of Technology Press, 1988.
[3] Y. Deng. Grey control system (second version). China University of Technology Press, 1985.
[4] Z. Mei. Grey absolute correlation degree and its calculation method. Systems Engineering, 1992 (5): 43-44.
[5] L. Wang and Y. Yang. A new method for calculating the grey relation degree and its application. Journal of Henan Agricultural University, 1996 (4): 401-403.
[6] W. Tang. T-type correlation degree and its calculation method. Mathematical Statistics and Management, 1995 (1): 34-37.
[7] Y. Chen and Y. Wei. A new method for calculating the grey correlation degree. Journal of Leshan Normal University, 2010, 25 (5): 14-18.
[8] H. Wen and A. Goodman. Relationship between urban land price and housing price: Evidence from 21 provincial capitals in China. Habitat International, 2013, 40(7): 9-17.
[9] Y. Zhang, X. Hua and L. Zhao. Exploring determinants of housing prices: A case study of Chinese experience in 1999–2010. Economic Modelling, 2012, 29(6): 2349-236.