

Research on the Relationship between Influencing Factors and House Prices Changes

Zixuan Lyu *

New College, University of Toronto, Toronto, Ontario, M5S 2E8, Canada

* Corresponding author: zixuan.lyu@mail.utoronto.ca

Abstract. In recent years, the change in house prices has also become a main factor affecting economic changes. There are lot of research shows that the relationship between economics and house prices. This study focuses on how house prices change and what factors can lead to the change in house prices. The data from the Kaggle website, containing property prices in Delhi. By using the linear regression model, calculate 12 factors which are Price, Area, Bedrooms, Bathrooms, stories, main road, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea, and furnishing status to study the degree and trend of impact on house prices changes. The results show that area, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea will have a significant positive impact on price, and furnishing status will have a significant negative impact on price. Bedrooms have no impact on price. This study opens a new perspective and giving society a new thinking angle for the study of housing prices changes.

Keywords: House prices, economic, linear regression model.

1. Introduction

After reform and opening up, China's economy began to revive. However, the economic development between regions is uncoordinated, this phenomenon is caused by a variety of objective and subjective factors. China has the largest population in the world, and housing issues have always been a focus of people's livelihood [1]. Tripathi considered that the housing market could be a vital factor that can affect the economy of the country, much economic research and economic policy have evolved around the role of housing prices for macroeconomic policy [2]. To research the uncoordinated economic development between regions in China, the change in housing prices could be a factor that cannot be ignored.

There are two main opinions on judging the direction of China's estate. One faction believes that China's real estate industry is a huge bubble, and another faction considers that house prices will skyrocket in the second half of 2009 [3]. Facts show that in 2010, housing prices in Beijing increased by 42%, in Guangzhou by 23%, and the national average increase was 9.9%. Chen et al. applied the Campbell and Shiler model to conduct an empirical analysis of the rent-to-rent ratio in China's real estate market, concluded that China's real estate bubble has existed for a long time [4].

Aimi demonstrated that China's real estate market at that time had some similarities with Japan's bubble economy [5]. Yu pointed out Real estate policy has a greater impact on housing prices than economic development [6]. Despite the Chinese government has carried out relevant macro-control measures to inhibit the phenomenon of skyrocketing housing prices, developers and home buyers continue to participate in the market and the prices cannot go down forever [7]. Therefore, it is necessary to explore the causes of changes in housing prices.

Population is the basis for national economic development and real estate. Wang considered that house prices have the mediating effect between the economic density and population density, that is, changes in the economic density will cause changes in housing prices then indirectly affect population density in this city [8]. Wu pointed out that housing type has a stable impact on out-of-town young talents' settlement intention. A reasonable house prices is a key factor affecting the willingness of young talents to settle in big cities [9]. Except for the external macro impact, there are still various determinants can affect the housing price. From 2000 to 2016, the average transaction price of various types of commercial housing in China continued to rise. In terms of categories, the prices of villas

and high-end apartments increased the most, followed by residential commercial housing. Price increases for other types of commercial housing were within 3 times [10]. By Liu's research, this paper studied two different types of houses: The townhouse and the apartment, to find the relationship between the house size, age and the house prices. The results show that, for townhouses, the greater space land makes the house prices lower, for apartments the higher housing age makes the house prices lower [3]. Jiang et al. used Eviews software to study housing prices in 31 key Chinese cities and revealed the reasons for high growth in housing prices from the perspectives of internal direct factors and external brief factors. The research shows that the land purchase price has a greater positive impact on the average selling price of residential commercial housing [7]. From Ebru's research, finding age, security, increasing of the number of bathrooms, etc. will increase the house prices. Additionally, they studied the house attributes, and the results showed that location is the most powerful factor in house prices, but the age of building and structure of the house are also should be considered in the change of house prices [11]. However, their research doesn't show how those factors affect the house prices. As factors inherent in the house itself, the house types and house structure are stable and important considerations when exploring the impact on house prices, caused they are not affected by external macroeconomics.

Summarizing the above research findings, this paper aimed to study changes in housing prices through the internal factors of the house itself. Focus on the structure of the house, the relationship between room size and the number of rooms on house price changes. With building a multiple linear regression model, this paper will use a dataset which contains the price of house in the Delhi region and find the relationship among the house prices with area, bathrooms, bedrooms, and so on. This study hopes can be useful for policymakers to make optimized policies for the real estate industry, and for developers and the public to consider comprehensive housing purchases.

2. Methods

2.1. Data Source

The data for this paper is collected from the Kaggle website, containing property prices in Delhi, complied by Gaurav and updated in 2023.

2.2. Variable Introduction

The data used in this paper has 545 samples and contains 13 variables (Price, Area, Bedrooms, Bathrooms, stories, mainroad, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea, and furnishingstatus). Since the variable mainroad is an invariant quantification, this paper uses other 12 variables to studies the relationship among Price and other 11 factors (table 1).

Table 1. Logogram and sample range of the dependent variable (Y) and the 11 factors X

Variable	Logogram	Sample range
Price	Y	1.75-13.3m
Area	X1	1650-16.2k
Bedrooms	X2	1-6
Bathrooms	X3	1-4
Stories	X4	1-4
Guestroom	X5	YES or NO
Basement	X6	YES or NO
Hotwaterheating	X7	YES or NO
Airconditioning	X8	YES or NO
Parking	X9	0-3
Prefarea	X10	YES or NO
Furnishingstatus	X11	YES or No

*Sample range: The range of each variable

*m: Million

Table 1 shows the range for each variable, this paper uses logogram to show variables for convenience of writing. The sample size of this data is 545.

2.3. Method Introduction

This paper uses a multiple linear regression model, with set price be the dependent variable (Y), and the 11 factors are the independent variables (X). For the variables like Guestroom, since their sample range is YES or No, this paper let 0 represent YES and 1 represent No. Then this study uses SPSSAU to analyze whether the dependent variable Price has an influence relationship with the 11 factors and find what the influence relationship is for each variable through two perspectives: the direction and the degree.

The multiple linear regression model is used for statistical analysis to estimate the relationship between two or more independent variables and one dependent variable. the formula can be seen that:

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_p x_p + \varepsilon \quad (1)$$

There Y means the predicted value of the dependent variable, α is the Y-intercept, which is the value of Y when all other parameters are set to 0. $\beta_1 x_1$ is the regression coefficient β_1 of the first independent variable x_1 , it also means the effect that increasing the value of independent variable has on the predicted Y. ε is the mode error, to show how much variation there is in our estimate of Y.

This paper first calculates R2 value to analyze the model fitting situation and uses VIF value to determine whether the model has the collinearity problems. Then to find the best-fit line for each independent variable, calculate t statistic of the overall model and p value. The next step is to calculate the p value of each regression coefficient and use p-value to show the relationship between the dependent variable and independent variables.

3. Results and Discussion

3.1. Correlation Coefficient Analysis

Firstly, analyze the model fitting situation, that is, analyze the model fitting situation through the R2 value, and analyze the VIF value and F value to determine the collinearity problems.

Table 2 shows the correlation coefficient analysis of this model, from Table 2, this study finds the R2 is 0.675, which means the model has 67.5% degree of fit with the variables, area, bedrooms, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea, furnishing status can explain 67.5% of the price changes. Since the VIF value of each element are less than 5, which means this model doesn't have collinearity. From the F test was performed on the model, it was found that the model passed the F test ($F=100.529$, $p=0.000<0.05$), which means that at least one of area, bedrooms, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea, and furnishingstatus will have an impact on price, this model is meaningful.

Table 2. Correlation coefficient analysis

Elements	Regression coefficients	95% CI	Collinearity Diagnosis	
			VIF	Tolerance
constant	-242.466**	-3244822.112 ~ -1604516.866	-	-
X1	0.0258**	211.457 ~ 305.551	1.271	0.787
X2	9.308	-49317.571 ~ 235493.109	1.347	0.742
X3	97.505**	770835.430 ~ 1179268.996	1.284	0.779
X4	47.749**	351874.212 ~ 603125.693	1.448	0.691
X5	30.941*	49040.505 ~ 569779.912	1.212	0.825
X6	36.526**	147393.095 ~ 583142.802	1.320	0.758
X7	86.694**	426080.556 ~ 1307806.906	1.039	0.962
X8	84.961**	635733.548 ~ 1063495.748	1.207	0.828
X9	30.056**	185680.337 ~ 415445.959	1.195	0.837
X10	69.089**	464023.659 ~ 917767.483	1.130	0.885
X11	-23.006**	-354059.951 ~ -106077.075	1.087	0.920
Sample size	545			
R 2	0.675			
AdjustR 2	0.668			
F value	F (11,533)=100.529,p=0.000			

*Regression coefficients: For the convenience of calculation, this paper multiplies the unit of the variable by the fourth power of ten.

3.2. Regression Model Results

From Figure 1, it shows the regression coefficient of 11 independent variables and the relationship between independent variables with dependent variable. Since the value of regression coefficient are too big to represent as a formula, this paper will multiply the unit of the variable by the fourth power of ten and take 4 decimal places. Through calculations, this study yields the final linear regression equation:

$$Y = -242.4669 + 0.0258 * x_1 + 9.3087 * x_2 + \dots + + 69.0895 * x_{10} - 23.0068 * x_{11} \quad (2)$$

Where Y represents the price, $x_1 \dots x_{11}$ represents the 11 independent variables: area, bedrooms, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea, furnishing status.

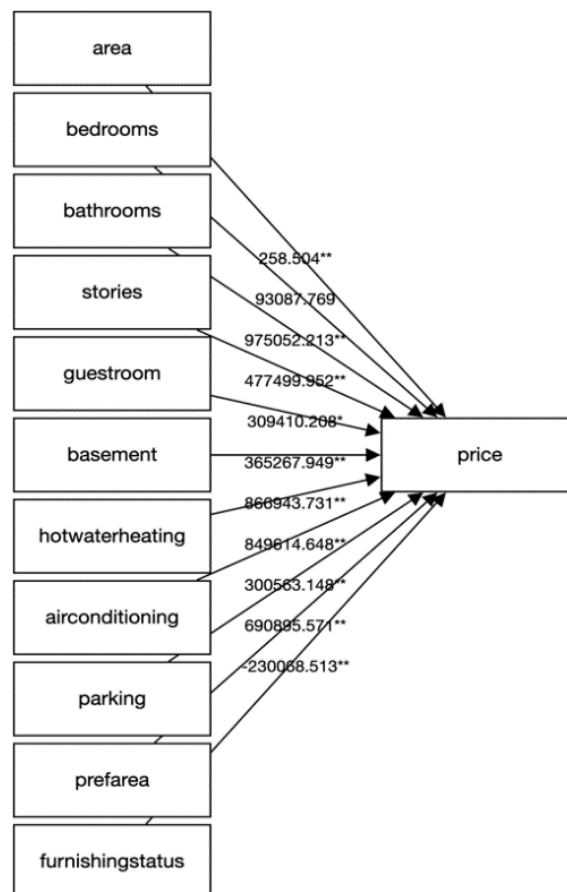


Figure 1. Linear Regression model schematic

Table 3. Linear regression model results

Elements	B	t	p
Constant	-242.4669489**	-5.794	0.000
X1	0.0258504**	10.769	0.000
X2	9.3087769	1.281	0.201
X3	97.5052213**	9.358	0.000
X4	47.7499952**	7.450	0.000
X5	30.9410208*	2.329	0.020
X6	36.5267949**	3.286	0.001
X7	86.6943731**	3.854	0.000
X8	84.9614648**	7.786	0.000
X9	30.0563148**	5.128	0.000
X10	69.0895571**	5.969	0.000
X11	-23.0068513**	-3.637	0.000
Sample size		545	

*B: Regression coefficient

**p value shows 0.000 means it is very small

In above Table 3, this study gets the conclusion from judging whether the p value is smaller than 0.05, which means the significance of X. Combine with B value can help to comparatively analyze the impact of independent variables X on dependent variable price Y. When the p value is smaller than 0.01 means this independent variable has a significant relationship with the dependent variable price. By following the result of this model, only the p value of bedrooms is 0.201, which is greater than 0.05, thus bedrooms do not affect price. Besides, other independent variables are all smaller than 0.05, that means area, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning,

parking, prefarea, and furnishing status will have an impact on price. The furnishing status's B value is negative shows it has significant negative effect on price. In other words, the other 9 independent variables which have positive B value will have positive impact on price. If the p value shows 0.000 ($0.000 < 0.01$) means this variable has highly significant in the model, this result also tells us that the p value of area, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea and furnishing status are all smaller than 0.01, which illustrate that these factors have a particularly strong relationship with price. Undoubtedly, through analyzing the result of the linear regression model, the 10 factors all have strong influence on price, prices change as these factors change, as this data has big sample size and large sample range, which means they have some reference and research value.

Compared to previous studies, they prefer to focus on what factors could affect house prices, there is no more detailed discussion of how factors will affect house prices. Most of studies tend to analyze from a macro perspective, such as combining economic or related policy changes to find the relationship of house prices. In constant, the factors covered in this study are more specific and comprehensive, this paper avoids possible errors from a single factor by increasing the sample size, passing the F test to determine collinearity problems, and further ensuring that the results of model analysis are accurate. This study also aimed to broaden the ideas of future research on house prices change and help developers and the public to consider comprehensive housing purchases.

4. Conclusion

The current research selected eleven factors that may be related to house prices changes, focusing on how they affect the price and the degree of influence, and concluded that area, bathrooms, stories, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea and furnishing status may have a significant impact on price, but the bedrooms have no impact on price.

Since the sample size cannot be infinitely large, and more potential factors may cause price changes that are not fully covered, although this study has tried our best to eliminate errors that may lead to the accuracy of the results, this study is not perfect, but still, there is meaningful and has a lot of value and advantage. On the one hand, the approach taken in this study is innovative. This paper uses graphics to intuitively analyze the relationship between the eleven related factors and house prices, making the research visual and making the results more concise, and easy to understand. The application of the multiple linear regression model has brought this study to a new height, opening a new perspective, and giving society a new thinking angle for the study of housing prices changes. On the other hand, housing issues are people's livelihood issues, maybe this paper should focus on "people" again and explore the reasons for changes in housing prices from the perspective of personal life needs. The 11 factors in this study are not comprehensive, just as the factor of the internal structure of the house, there are still lots of other factors that may affect house prices like community services, surrounding environment, etc. Whether these factors are associated with house prices requires further study, this paper hopes that those research results can provide some new ideas so that society and relevant government agencies can study housing prices changes in a more "humane" manner and improve people's quality of life.

References

- [1] Cheng Qiyun, Sun Caixin, Zhang Xiaoxing, et al. Short-term load forecasting model and method for power system based on complementation of neural network and fuzzy logic. Transactions of China Electrotechnical Society, 2004, 19 (10): 53 - 58.
- [2] Tripathi S. Macroeconomic determinants of housing prices: A cross country level analysis. Munich Personal RePEc Archive, 2019.
- [3] Liu Taicheng. A study on the determinants of housing price: Evidences from metropolitan area of Taichung. Networked Digital Library of Theses and Dissertations.

- [4] Zhao Zhijun. Analysis of China's real estate market characteristics and factors affecting housing prices. Institute of Economics, Chinese Academy of Social Sciences, 2019, 5 - 9.
- [5] Li Boning, Liao Wenchi, Sun Weizeng, et al. The US-China trade war: quantify the negative shocks to local housing markets & land-based finance in Chinese cities, 2023, 23 (17): 2 - 4.
- [6] Yu Huayi. China's house price: affected by economic fundamentals or real estate policy. Higher Education Press and Springer-Verlag, 2010, 5 (1): 25 - 51.
- [7] Jiang Yuxi, Qiu Lilan. Empirical study on the influencing factors of housing price –Based on cross-section data of 31 provinces and cities in China. Procedia Computer Science, 2022, 199: 1498 - 1504.
- [8] Wang Tianying. Urban Economic Density and Population Density: The Mesomeric Effect of Housing Prices. World Economic Exploration, 2023, 12 (3): 275 - 283.
- [9] Wu Kaize. Living in peace and working in contentment: on factors influencing migrant young talents' settlement intention in metropolises. Journal of East China Normal University (Philosophy and Social Sciences), 2023, 55 (4): 162 - 175.
- [10] Ebru C, Eban A. Determinants of house prices in Istanbul: a quantile regression approach. Quality & quantity, 2011, 45 (2): 305 – 317.
- [11] Owusu-Manu D G, et al. Housing attributes and relative house prices in Ghana. International Journal of Building Pathology and Adaptation, 2019, 37 (5): 733 - 746.