Proceedings of the 3rd International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/2025.23756

The Research on Factors Influencing Housing Prices-Take Chicago as an Example

Zhaoyi Chen[1*], Siqian Ji[2], Ziyu Li[3]

[1]Capital Normal University High School, Beijing, China
[2]School of Adocte, Shanghai, China
[3]The Second High School attached to Normal University, Beijing, China
**Corresponding Author. Email: ccczy080325@gmail.com*

Abstract: This paper purposes to research what factor is influenced the house price in Chicago. The Multiple Linear Regression will be used as the main method in the article. The 158 samples are chosen to analyze the factors. According to this data, there are 8 variables will be related with housing price. There are the number of bedrooms, spaces, rooms, lots, taxes, bathrooms, garages, conditions respectively. Also, this paper considers the area, stories, guestroom, basement, hotwaterheating, air-conditioning, parking, prefarea, furnishing status, and use Pearson correlation to show the strength of them. To prove the accuracy of the set of data, the model will use the value of F to check. The fact is verifying that area, bedrooms, bathroom have significant relation with the house price in Chicago and the other factors are not have direct relationship with price. Overall, the house price in Chicago can be influenced by the condition of the environment around the house.

Keywords: Housing prices, Chicago, multiple linear regression, influencing factors, correlation analysis.

1. Introduction

This study uses XGBoost and SHAP value methods to analyze panel data of 70 Chinese cities from 2009 - 2019, aiming to identify key factors behind such price differentials. Findings show that expectation factors drive housing price divergence. Monetary, demand and supply factors aren't primary causes, yet can influence prices by shaping expectations. Thus, the essence of "city - specific policies" is stabilizing public price expectations. First - tier cities should control land costs and boost supply to counter upward price expectations. Second - and third - tier cities need to improve infrastructure and living standards to meet housing demand and prevent downward price expectations due to population outflow [1].

By reviewing the research findings of scholars on factors influencing housing prices, this study summarizes seven main factors potentially affecting Beijing's commercial housing prices from supply, demand, and macro - economic perspectives. These factors include the sales area of commercial housing, permanent population, the proportion of the population aged 15-64, per capita GDP, consumer price index, the land area purchased by real estate development enterprises, and inflation rate. Using Stata, a multiple linear regression model is established to analyze these related variables. The results show that all these factors have a significant impact on Beijing's housing prices. Based on the regression results, three policy recommendations are put forward for the government's macro -

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.23756

control: encouraging mid-to-low-end housing projects for rigid demand, cracking down on "speculative housing" phenomena, and vigorously developing the housing rental market and affordable housing [2].

The research covers housing prices and determinants in 30 Chinese regions (excluding Tibet, Hong Kong, Macau, and Taiwan), splitting the national market into the East, Central, and West. Drawing on prior work, 12 factors from supply, demand, policy, and other aspects are identified as key drivers of housing prices both nationally and regionally. To verify these factors, a panel data grey relational analysis is created. It assesses factor - price correlations, weeding out weak - impact factors, and conducts descriptive statistics on housing prices and related elements. Through multiple statistical tests, an individual - time double - fixed - effects panel data model is set up for housing prices and their determinants across China and in the three regions. The maximum likelihood method is used for model estimation. Accounting for spatial effects, a spatial Durbin model is crafted for national and eastern housing markets, and a spatial lag model for central and western ones, based on spatio-temporal price analysis and spatial tests. Spatial interactions among housing prices and their determinants are explored too. In sum, there are notable regional housing price disparities in China. Based on model analysis, specific conclusions and suggestions are proposed for housing price development in the three major regions [3].

Housing is fundamental for people's livelihood, and housing price trends draw significant public attention. With the growing share of second - hand housing transactions in the real estate market, accurately predicting second - hand housing prices has become crucial for buyers, sellers, and real estate agents. This study aims to construct a price evaluation model to facilitate more accurate transactions. Most existing research focuses on predicting average regional real estate prices from aspects like population, macro - policies, and regional economy. Fewer studies build housing price evaluation models based on property - specific attributes. This research takes Beijing's second - hand housing as an example, using 13,260 transaction records from Mark Data Network from 2020 - 2021. It considers both hard construction features (16 indicators such as location, layout, area) and soft attention-related indicators (views, follows, price adjustments, and transaction cycle). Grey relational analysis determines the correlation coefficients between these factors and housing prices. Copula function reveals the influence mechanism. The results show that floor area, location, building type, views, and transaction cycle are the top five influential factors with asymmetric tail correlations with housing prices. These factors are then input into a neural network and multi - objective algorithm to build an optimized prediction model (PCFM). PCFM achieves a prediction accuracy of 1.4%, outperforming other models. Finally, suggestions are provided for policymakers, real estate agents, and homebuyers to promote the healthy development of the real estate industry [4].

This article provides the trend of the house price in Chicago in 2024 and express the vision of the future, analysis the factor of house price [5]. By research the community with the highest increase in house prices in 2000, the factors will be obtained such as geographical position, supporting facilities [6]. Not only that, for better exploration that the data is estimated based on sales prices and appraised values, with the first quarter of 1995 as the base period of 100, and it is updated quarterly. It can serve as an important reference for studying the trend of housing price changes in Chicago [7].

In addition to studies on the Chinese housing market, research on Australian housing prices has also provided useful insights. For example, a study in Greater Sydney found that the quality and proximity of public schools significantly impact housing prices, with primary schools having a greater effect than high schools [8]. Another study used data science methods to analyze housing prices, incorporating both micro-level and macro-level data. The HNED model, which includes housing attributes, neighborhood characteristics, and demographic features, showed strong explanatory power with an R^2 value of 0.88 [9]. Research on the causes of rising housing prices in Australia identified key factors such as interest rates, investment demand, and population growth [10].

DOI: 10.54254/2753-8818/2025.23756

Specifically, this study will focus on analyzing the factors influencing housing prices in Chicago using multiple linear regression analysis. The goal is to identify the primary drivers of price fluctuations and quantify their impact.

2. Methods

2.1. Data source

The dataset used in the article is obtained from Kaggle website that is describe house price in Chicago. It was collected by Tawfik. This dataset selected 158 of all data they research as sample. The original dataset is no change in .csv format.

2.2. Basic information of variables

This study's variables are meticulously selected to comprehensively capture aspects influencing Chicago housing prices (Table 1). For instance, area directly reflects the size of a house, which typically correlates with price. Bedrooms and bathrooms are key functional elements; more of each can enhance a home's usability and value. Variables like guestroom, basement, and air - conditioning represent additional amenities. The binary variables, such as whether there is a guest room or basement, simplify data analysis by indicating the presence or absence of these features. Understanding these variables is crucial for accurately interpreting the subsequent regression analysis results.

Table 1: Basic information of variables

Variable Name	Variable Definition	Value Range
area	The area of the house	1650-16200
bedrooms	The number of bedrooms	1-6
bathrooms	The number of bathrooms	1-4
stories	The number of floors of the house	1-4
guestroom	Whether there is a guest room	0 (no) or 1 (yes)
basement	Whether there is a basement	0 (no) or 1 (yes)
Hot water heating	Whether there is hot water heating	0 (no) or 1 (yes)
furnishing status	The furnishing condition of the house	0 (unfinished) or 1 (furnished)
prefarea	Whether it is in a prime area	0 (no) or 1 (yes)
parking	The number of parking spaces	0-3
air-conditioning	Whether there is air conditioning	0 (no) or 1 (yes)

2.3. Method introduction

The paper uses Multiple Linear Regression model to analyze key factors affecting housing prices. The research objective mainly focuses on identifying which variable most strongly affect pricing and finding out the most optimized model that measures how much each influencing factor contributes to price differences.

The Multiple Linear Regression model helps determine the linear correlation between one dependent variable and two or more independent variables. The model is defined as:

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\ldots+\beta_n x_n+\epsilon \quad (1)$$

DOI: 10.54254/2753-8818/2025.23756

where y stands for the dependent variable (housing price), β_0 the intercept, $\beta_1, \beta_2 \ldots \beta_n$ are the coefficients for the independent variables, $x_1, x_2 \ldots x_n$ are the independent variables (factors influencing housing prices), and ϵ is the error term.

3. Results and discussion

3.1. Correlation analysis

The relevance analysis evaluates the relationships between housing prices and 11 potential independent variables as shown in the table 2.

Table 2: Correlation analysis between dependent and independent variables

	price
area	0.242*
bedrooms	0.304**
bathrooms	0.319**
stories	0.017
guestroom	0.031
basement	0.093
hotwaterheating	-0.025
furnishingstatus	-0.069
prefarea	0.107
parking	0.411**
airconditioning	0.098

* $p<0.05$ ** $p<0.01$

Based on the table 2, there are some key findings. First, bathrooms, bedrooms, and parking displayed moderate-to-strong positive correlations with price. Second, area showed a weak but statistically significant positive correlation. Third, variables such as stories, basement, and air-conditioning ($r = 0.098$) exhibited approximately zero correlations with price ($p > 0.05$).

As for the analysis of inter-variable (figure 1 shown), the correlations among potential influencing factors were uniformly weak ($|r|$ approximately equal to zero), with no statistically significant relationships. The computed correlation coefficients (0.129, 0.073, -0.172, -0.140, 0.090, -0.139, -0.122, 0.181, 0.095, -0.093 respectively) all demonstrated negligible magnitudes approaching zero. Furthermore, all corresponding p-values exceeded the 0.05 significance threshold. These statistical outcomes collectively indicate the absence of meaningful associations between the area measurement and the other investigated housing characteristics in this study.

DOI: 10.54254/2753-8818/2025.23756

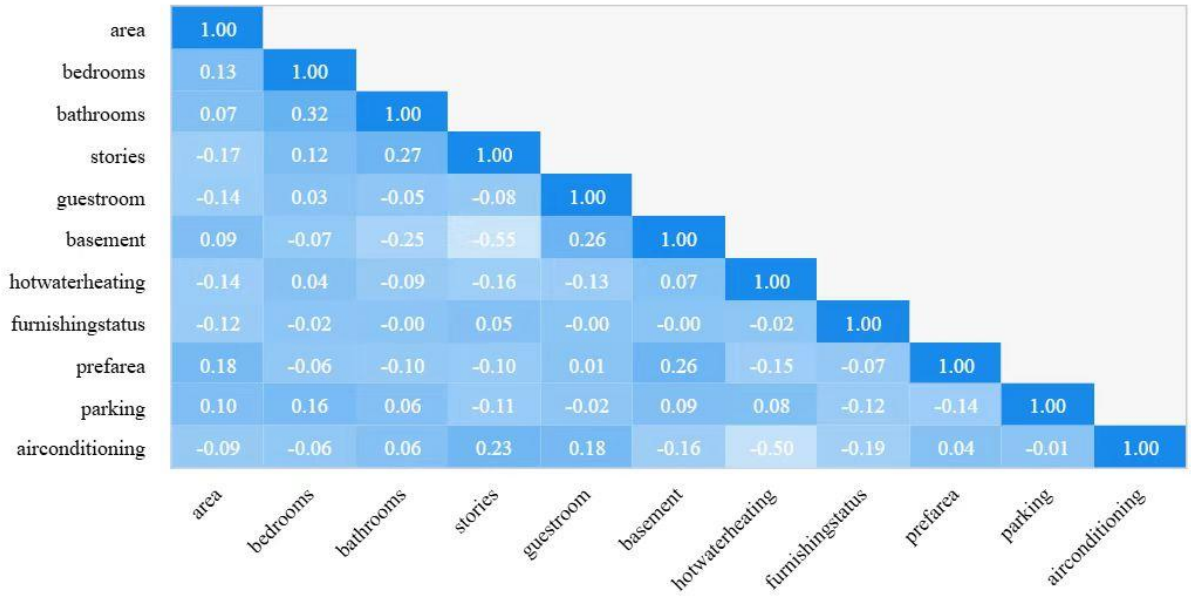


Figure 1: Relevance analysis among independent variables

3.2. Multiple linear regression analysis

As the result shown in the table 3, the multiple linear regression model for this dataset is:

$$y = 732673.319 + 124.558x_1 + 650566.462x_2 + \cdots + 78390.114x_{11} \quad (2)$$

The regression model (F = 4.652, p < 0.001) explained 36.8% of the variance in housing prices (R^2 = 0.368), indicating that the influencing factors collectively account for a moderate portion of price fluctuations. Two variables demonstrated statistically significant effects: bathrooms and parking, reflecting buyers' prioritization of functional living spaces and vehicular accessibility. The positive coefficients suggest that each additional bathroom increases prices by approximately 650566 units, while an extra parking space adds 660661 units.

Moreover, variables initially showing bivariate correlations with price, including area and bedrooms, became non-significant in the model. Non-significant results were also observed for stories, air-conditioning, and furnishing status. Variance Inflation Factor (VIF) values for all influential factors ranged between 1 and 5, confirming no severe multicollinearity.

Table 3: Results of multiple linear regression analysis

	Non-standardized coefficients		Standardized coefficients	t	p	Collinearity Dianosis	
	B	Standard Error	Beta			VIF	Tolerance
Constant	732673.319	1664938.869	-	0.440	0.661	-	-
area	124.558	65.678	0.177	1.897	0.061	1.219	0.820
bathrooms	650566.462	230693.362	0.265	2.820	0.006**	1.227	0.815
stories	87504.025	144525.352	0.065	0.605	0.546	1.606	0.623
bedrooms	372422.350	220797.720	0.155	1.687	0.095	1.182	0.846
guestroom	81401.929	304786.009	0.025	0.267	0.790	1.191	0.839
basement	403708.567	333873.702	0.134	1.209	0.230	1.702	0.588
hotwaterheating	486872.335	510848.994	0.099	0.953	0.343	1.494	0.669

DOI: 10.54254/2753-8818/2025.23756

Table 3: (continued).

airconditioning	530970.100	339127.011	0.166	1.566	0.121	1.565	0.639
parking	660661.380	159352.363	0.370	4.146	0.000**	1.108	0.903
prefarea	434783.946	276309.617	0.145	1.574	0.119	1.183	0.845
furnishingstatus	78390.114	172544.131	0.041	0.454	0.651	1.113	0.899
R^2			0.368				
Adjusted R^2			0.289				
F			F (11,88) =4.652, p=0.000				
D-W value			0.797				

4. Conclusion

This study utilized the Multiple Linear Regression model to explore the key factors influencing housing prices in Chicago. Through relevance analysis, it was found that bathrooms, bedrooms, and parking had moderate - to - strong positive correlations with housing prices. This indicates that the number of functional spaces like bathrooms and bedrooms, as well as the availability of parking, are important considerations for homebuyers in Chicago, and these factors can significantly affect housing prices. The area also showed a weak but significant positive correlation, suggesting that larger houses generally tend to have higher prices. However, variables such as stories, basements, and air - conditioning had nearly zero correlations with housing prices, meaning they may not be primary determinants of housing prices in this context.

The multiple linear regression analysis showed that the model could explain 36.8% of the variance in housing prices. Bathrooms and parking had significant effects, with each additional bathroom increasing prices by approximately 650,566 units and an extra parking space adding 660,661 units. Nevertheless, some variables that initially showed correlations with housing prices, like area and bedrooms, became insignificant in the model. This might be due to the complex interactions among various factors or the limitations of the model.

Overall, this research identified some important factors affecting housing prices in Chicago. However, it should be noted that the model has its limitations. The relatively small sample size of 158 samples may not fully represent the entire Chicago housing market, which could lead to potential biases in the results. Also, there may be other unconsidered factors, such as local economic development trends, changes in interest rates, and new housing policies, that could impact housing prices. Future studies could expand the sample size, incorporate more comprehensive data, and consider additional influencing factors to improve the model's accuracy and explanatory power. This would provide a more in - depth understanding of the housing price mechanisms in Chicago and offer more reliable references for relevant stakeholders.

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] *Chen, X.L., Chen, K., Wang, Z.R. and Xiao, Z.Y. (2024) Research on the influencing factors of housing price differentiation among cities. Economic Theory and Business Management, 2, 49-64.*

[2] *Meng, K.Y. (2023) Analysis of the influencing factors of housing prices and policy recommendations-Taking Beijing as an example. Economic Research Guide, 22, 32-34.*

[3] *Zha, H. (2023) Analysis of the influencing factors of housing prices in China based on panel data. Master's Thesis, Nanjing Audit University, Master.*

DOI: 10.54254/2753-8818/2025.23756

[4] *Yan, L.Y. (2023) Prediction of second - hand housing prices in Beijing and analysis of influencing factors based on optimization algorithms. Master's Thesis, Dongbei University of Finance and Economics, Master.*
[5] *Norada Real Estate Investments. (2025) Chicago Housing Market: Trends and Forecast 2025-2026. Norada Real Estate Investments. Retrieved from https://www.noradarealestate.com/blog/chicago-real-estate-market/*
[6] *Federal Reserve Bank of Chicago. (2018) Determinants of Housing Values and Variations in Home Prices Across*
[7] *Neighborhoods in Cook County. Federal Reserve Bank of Chicago. Retrieved from https://www.chicagofed.org/publications/profitwise-news-and-views/2018/determinants-of-housing-values-and-variations-in-home-prices-across-neighborhoods-in-cook-county*
[8] *Lu, Y. Shi, V. and Pettit, C.J. (2023) The Impacts of Public Schools on Housing Prices of Residential Properties: A Case Study of Greater Sydney, Australia. ISPRS Int. J. Geo-Inf., 12, 298.*
[9] *Kou, J.Y. (2022) Analysing Housing Price in Australia with Data Science Methods. PhD thesis, Victoria University.*
[10] *Rahman, M.M. (2010) The Australian housing market – understanding the causes and effects of rising prices. Policy Studies, 31(5), 577-590.*