# Research on the influence of metro on housing price via difference-in-difference model

Wensong Chen, Yongmei Ding*, Yue Dong
College of Science, Wuhan University of Science and Technology, Wuhan 430065, Hubei, China

## ABSTRACT

In order to explore the impact of urban rail transit on the second-hand real estate market, based on the difference-in-difference model, this paper takes Wuhan Metro Line 7 as the research object, and makes an empirical analysis from two perspectives of different districts and different distance ranges. The results show that the operation of Wuhan Metro Line 7 has a statistical significance on the housing price in Jiangxia District, Hongshan District, Dongxihu District and Jiang'an District, and the impact coefficients are 1426.2, 786.7, 647.3 and 520.2, respectively. In terms of the different distances of Metro Line 7, the opening of the subway within 1200 meters has a significant positive influence on the second-hand housing prices along the periphery, and the degree of impact is 500m, 800m-1200m, and 500m-800m in descending order. Quantitative analysis of the impact of the subway on the surrounding housing prices can provide the government with a reasonable guide resident to purchase houses and provide relevant evidence for the scientific implementation of land acquisition and demolition.

**Keyword:** Difference-in-difference model, Wuhan metro line 7, second-hand housing prices

## 1. INTRODUCTION

With the increase of urban population density, housing shortage, traffic congestion and a series of problems are increasingly prominent. With its significant advantages of safety, comfort, high efficiency and environmental protection, urban rail transit alleviates the traffic pressure on ground roads to a great extent and reduces the travel cost of residents. However, while improving the traffic conditions, urban rail transit will also increase the functional utility of schools, shopping malls and other places, thus driving the rise of housing prices in areas along the route. When people consider the location of a house, it will be taken into account as a factor affecting the price of a house. Therefore, it is of great practical significance to study the impact of subway on second-hand real estate prices, which is conducive to real estate developers' more reasonable pricing, relevant policy departments' scientific evaluation of land value, and buyers and renters make rational housing choices.

Since the opening of the world's first city subway in 1863, scholars at home and abroad have conducted relevant studies on the relationship between subway and real estate. Based on the hedonic model, Du[1] analyzed the influence of Shijiazhuang Metro Line 1 on housing prices along the line, and found that the value added effect on housing prices was the best within the range of 700-900m. Zhang[2] took Xi'an Metro Line 2 as the research object, and found that there was a negative correlation between the housing price and the distance to the subway station within the range of 150-1200m. Taking Nanchang Metro Line 1 as an example, Wang[3] believed that within 1500m, the housing price decreases with the increase of the distance from the subway station. Through the analysis of the influencing factors of the housing price, it was found that the opening of rail transit has a significant impact on the housing price[4-7].

Based on the hedonic price theory, Mohammad[8] analyzed that the maximum effect of Dubai subway on residential prices is within 701-900m of the metro station. Seo[9] believed that the opening of the Seoul Metro has influenced people of different economic classes to choose different apartment types. Taking the Toronto subway as an example, Bae[10] showed that the distance between a house and the subway station would have a significant impact on housing prices before the opening of Seoul Metro Line 5. Considering the space-time effect comprehensively, it is found that the subway has different influence on the housing price in different construction periods, and the distance from the subway station to the house has different influence on the housing price[11, 12].

*dingyongmei@wust.edu.cn

International Conference on Computer Application and Information Security (ICCAIS 2021),
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12260, 1226020 · © The Authors.
Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.2637401

Analysis on the influence of urban rail transit on housing price, the most of existing research using hedonic model, the hedonic price theory, which reflect the contribution degree of various characteristics to the price. But it is difficult to ensure that all the important variables affecting prices are taken into account, Furthermore, time-induced changes are difficult to eliminate, and easy to cause multicollinearity problems. The difference-in-difference model is used in this paper, introducing time dummy variable, the price change and factor change caused by time are well separated out, and the paper reflect the impact of subway operation more clearly from the perspectives of different districts and different ranges, greatly enhance the accuracy of the model results, it can systematically analyze the effect mechanism of urban rail transit on housing price.

## 2. RESEARCH METHOD

### 2.1. Introduction of DID model

Difference-In-Difference (DID) model, also known as multiple Difference method, is a kind of evaluation model mainly used for policy or project effectiveness. Compared with the traditional regression model, DID model can effectively analyze the policy impact by combining the "before and after difference" with the "with or without difference". The core idea of the model is to construct a double estimator, which can be obtained by combining the comparison before and after the intervention of a policy or project and the comparison between the control group and the treat group. In other words, the difference between the treat group and the control group before and after the policy impact is the net effect of the policy.

$$d_{ID} = \Delta\overline{Y}_{treat} - \Delta\overline{Y}_{control} = \left(\overline{Y}_{treat,t1} - \overline{Y}_{treat,t0}\right) - \left(\overline{Y}_{control,t1} - \overline{Y}_{control,t0}\right) \quad (1)$$

where, $d$ is the double differential estimator, $Y$ is the research variable, $\overline{Y}$ represents the mean value of the research variable, the footnote treat represents the experimental group or treatment group, control represents the control group or control group, $t_0, t_1$ respectively represent before and after the policy intervention.

In order to judge the difference before and after a policy intervention and satisfy the requirement that the observation points are identical distribution when sampling the population in different time periods, DID model introduced the time dummy variable. To judge whether the policy has different effects on different groups, the group dummy variable was added to group the research objects. Finally, in order to satisfy the premise hypothesis of DID model, the following hypotheses should be satisfied:

Hypothesis 1: The opening of Wuhan Metro Line 7 has no influence on the areas it has not passed through.

Hypothesis 2: Except for the influence of the opening of Metro Line 7, the external influence on the treat group community is the same as that of the control group community.

Hypothesis 3: The distribution of important features between the treat group and the control group is stable and does not change before and after the opening of Metro Line 7.

### 2.2. Independent mixed cross-section data DID model

The independent mixed cross-section data refers to the results of random sampling from the population at different time points, which maintains the independence of the samples. Since this kind of data is not controlled by time nodes, more sample capacity can be acquired, and better fitting effect and more accurate estimators can be obtained. The basic form is as follows:

$$y = \beta_0 + \beta_1 T + \beta_2 W + \beta_3 T \cdot W + \varepsilon \quad (2)$$

where, $y$ represents the explained variable, $\beta_0$ represents the intercept term, $\beta_1, \beta_2, \beta_3$ are regression coefficients, $W$ represents the group dummy variable, W=1 indicates that the sample belongs to the treat group, W=0 indicates that the sample belongs to the control group.$T$ is the time dummy variable, $T = 0$ means the sample is before the occurrence of the policy, $T = 1$ means the sample is after the occurrence of the policy, and $\varepsilon$ is the error term. $\beta_0$ represents the characteristics of the control group before the policy occurred, $\beta_1$ represents the change of the control group before and after the policy occurred, $\beta_2$ represents the difference between the treat group and the control group before the policy occurred, and $\beta_3$ represents the effect of the policy intervention.

### 2.3. DID model of panel data

The data studied in this paper are the second-hand housing prices of residential areas along Metro Line 7 in July 2018 and December 2019 at two-time nodes, which are panel data of two periods, and have the characteristics of cross-section and time sequence. Data collection is required for the same individual at different time points. Different from the independent mixed cross-section data, Panel data is not independently distributed at different time points. Since it is impossible to guarantee that individual factors are completely unrelated to grouping variables, we need to separate these factors and call them as a. Grouping variable W and time dummy variable T were introduced respectively, and the basic form of DID model satisfied was:

$$y = \beta_0 + \beta_1 T + \beta_2 W + \beta_3 T \cdot W + a + \varepsilon \tag{3}$$

Since a is unobservable, information cannot be collected.a is removed by the first-order difference method, and the following expression is obtained:

$$\Delta y = \beta_1 + \beta_3 T \cdot W + \varepsilon' \tag{4}$$

where, $\varepsilon$, $\varepsilon'$are all error terms. $\beta_3$ represents the interaction between the time variable and the sample grouping variable. The magnitude and direction of $\beta_3$ reflect the net effect of the policy of double difference, that is, the impact of the opening of Metro Line 7 in this paper.

## 3. EMPIRICAL ANALYSIS

### 3.1. Research objects and data description

Wuhan Metro Line 7 is the eighth completed and operational subway Line in Wuhan, Hubei Province. occupying the highest speed, the highest passenger capacity and the highest transfer of Wuhan subway. As of December 2018, Wuhan Metro Line 7, has been opened to traffic. It crosses six districts and starts from North Yuanboyuan Station in Dongxihu District in the north, passes through Jianghan District, Jiang'an District, Wuchang District and Hongshan District, and ends at Qinglongshan Metro Town Station in Jiangxia District in the south, with 26 stations in total. As a result, the opening of Line 7 not only improves the citizens' happiness of life, but also drives the development along the line. It lays a good foundation for researching the influence range of Wuhan Metro Line 7 in different areas.

We take Wuhan Metro Line 7 as the main axis, and the six districts it passes through (as the treat group) and the unpassed district (as the control group) as the research area. The research unit is within 1200m from the subway station, which is divided into three sections: 12-500m, 500-800m and 800-1200m. The second-hand housing price in these areas is taken as the research object. A total of 128 residential samples were selected for the study, and the average price of second-hand housing in July 2018 and December 2019 in the sample residential areas were collected. The relevant data came from Fangtianxia, Anjuke and Baidu Map.

### 3.2. The impact of Wuhan Line 7 on housing prices in districts along the line

We collected the sample plot in its regional grouping, The time dummy variable (0, before opening; 1, after opening), the group dummy variable (1, treat group; 0, control group), interactive dummy variable in time and group (1, opening and the treat group; 0, otherwise) were introduced, after first order difference of OLS regression for each district, The regression results are summarized in Table 1.

According to the coefficient symbol of $\beta_3$ estimate and the corresponding P value in Table 1, firstly, the opening of Metro Line 7 has a significant positive influence on the Dongxihu District, Jiang'an District, Hongshan District, and Jiangxia District. The order of influence degree is as follows: Jiangxia District, Hongshan District, Dongxihu District and Jiang'an District. Since Metro Line 7 is the first subway to Hongshan District and Jiangxia District, and the average price of second-hand houses in Jiangxia District is lower than that in other districts, therefore, Jiangxia District and Hongshan District had the greatest influence after the opening of the subway. Secondly, the opening of Line 7 has a positive impact on Jianghan District and Wuchang District, but it is not significant, mainly because before the opening of Line 7, there was a subway operation in these areas, the transportation system is relatively developed, they were prosperous areas, so the overall housing price level is relatively high, and the opening of Line 7 has no significant impact.

Table 1. DID model results of each district.

| District | Coefficient | Estimate | Std. Error | *t* value | Pr (>\|t\|) | | Significance |
|---|---|---|---|---|---|---|---|
| Dongxihu District | $\beta_1$ | -1498.9 | 188.7 | -7.942 | 2.43e-09 | *** | Significant |
| | $\beta_3$ | 647.3 | 318.4 | 2.033 | 0.0497 | * | Significant |
| Jianghan District | $\beta_1$ | -1413.60 | 144.50 | -9.783 | 6.26e-12 | *** | Significant |
| | $\beta_3$ | 79.53 | 235.97 | 0.337 | 0.738 | | No Significant |
| Jiang'an District | $\beta_1$ | -1413.6 | 206.4 | -6.850 | 1.53e-08 | *** | Significant |
| | $\beta_3$ | 520.2 | 298.1 | 1.745 | 0.0877 | . | Significant |
| Wuchang District | $\beta_1$ | -1413.6 | 174.4 | -8.106 | 8.66e-11 | *** | Significant |
| | $\beta_3$ | 192.0 | 238.0 | 0.807 | 0.423 | | No Significant |
| Hongshan District | $\beta_1$ | -1413.6 | 179.4 | -7.879 | 5.45e-09 | *** | Significant |
| | $\beta_3$ | 786.7 | 348.7 | 2.256 | 0.031 | * | Significant |
| Jiangxia District | $\beta_1$ | -1413.6 | 136.3 | -10.371 | 1.69e-12 | *** | Significant |
| | $\beta_3$ | 1426.2 | 227.5 | 6.269 | 2.72e-07 | *** | Significant |

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

### 3.3. The impact of Wuhan Line 7 on housing prices in different distances along the line

Through the research in the last section, we analyzed that the opening of Metro Line 7 had no significant impact on the housing price of Jianghan District and Wuchang District. Therefore, we eliminated the data of these two districts and analyzed the samples according to the three interval segments divided above. The time dummy variable (0, before opening; 1, after opening), the group dummy variable (1, treat group; 0, control group), interactive dummy variable in time and group (1, opening and the treat group; 0, otherwise) were introduced. After first order difference of OLS regression for each district, the regression results are summarized in Table 2.

Table 2. DID model results in each distance interval.

| Distance | Coefficient | Estimate | Std. Error | *t* value | Pr(>\|t\|) | | Significance |
|---|---|---|---|---|---|---|---|
| < 500m | $\beta_1$ | -1498.9 | 188.3 | -7.958 | 2.32e-09 | *** | Significant |
| | $\beta_3$ | 1001.3 | 317.8 | 3.151 | 0.00332 | ** | Significant |
| 500m-800m | $\beta_1$ | -1413.6 | 204.9 | -6.898 | 1.81e-08 | *** | Significant |
| | $\beta_3$ | 699.4 | 307.4 | 2.275 | 0.0279 | * | Significant |
| 800m-1200m | $\beta_1$ | -1413.6 | 182.4 | -7.748 | 4.64e-10 | *** | Significant |
| | $\beta_3$ | 785.5 | 255.5 | 3.074 | 0.00345 | ** | Significant |

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In Table 2, the coefficient symbols of $\beta_3$ are all positive, and the corresponding *P* values are all less than 0.05, indicating that the opening of Metro Line 7 within 1200m has a significant positive impact on the second-hand housing prices along the periphery, meaning that the opening of Metro Line 7 will promote the rise of the housing prices within this range.

From the size of $\beta_3$ coefficient, the order of the influence degree from large to small interval is within 500m, 800m-1200m, 500m-800m. It is obvious that the effect does not decrease with the increase of distance. The main reason is that the economic development level of each district is not consistent, and the data of the treat group of 800-1200m in this study are mostly from Jiangxia District and Hongshan District. We have analyzed that the opening of subway has the greatest impact on these two districts, so that the housing price does not necessarily decline with the increase of distance within the range of 1200m along the line. It shows that the subway operation has different influences on different districts, which leads to different influences on different distance ranges.

### 3.4. Comparison of influence effects

This section mainly discusses the impact of time and the opening of the subway on different districts and different distances along the line. After data standardization, OLS regression results are as follows:

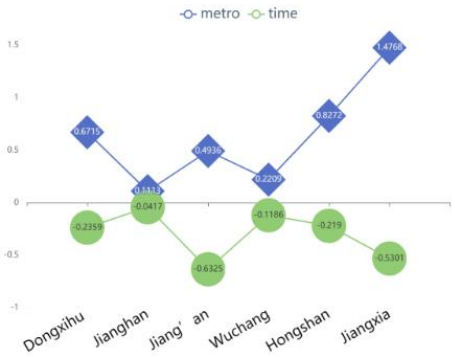


Figure 1. Influence of time and metro opening on different districts.

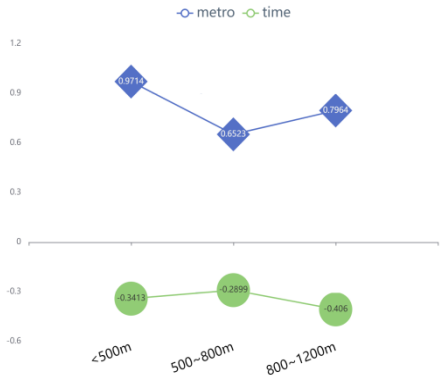


Figure 2. Influence of time and metro opening on different distance intervals.

Figures 1 and 2 show the influences of time and the opening of subway on different districts and distances. The opening of subway has a positive effect on different districts and different distances, whereas the time has a negative effect. On the whole, the degree of positive impact is significantly greater than the degree of negative impact, which further indicates that the opening of the subway has a significant positive effect on the second-hand housing prices along the line.

## 4. CONCLUSION AND RECOMMENDATION

This paper adopted two-time nodes of Wuhan Metro Line 7 in the first half year and one year after its opening to traffic, and analyzed the different districts the subway passes through and the different distance ranges around the subway. The empirical results show that:

(1) The opening of Metro Line 7 has a positive impact on the second-hand housing prices in different districts along the line, and the impact degree is from large to small: Jiangxia District, Hongshan District, Dongxihu District, Jiang'an District, Wuchang District and Jianghan District, among which the impact on Wuchang District and Jianghan District is not significant.

(2) The opening of the subway has a positive influence on the second-hand house prices in different distances, with the largest influence on the range within 500m, followed by 800-1200m, and the least impact on the range between 500-800m.

Based on the above research, by grasping the relationship between the opening of subway and the housing price along the line, the value of land along the line can be improved, the economic development along the line can be promoted, and the reference basis for the relevant decision makers can be provided. For the government departments, using the existing subway system and planning the subway extension line, it does not need to spend a huge amount of money on the whole subway line to drive the suburban economic development effect. For houses buyers, the opening of the subway does have a positive impact on the housing price. Compared with the city center, the second-hand houses farther away from the city center are more likely to appreciate in value. The second-hand houses along the subway are also a good choice

compared with other areas. For investors, they can invest in advance during the construction of the subway, reduce the cost input, and develop the commercial circle along the line in a targeted way, so as to seek greater benefits.

## REFERENCES


[1] Du, Q. J. and Wang, W. X., "Analysis on the influence of urban metro construction on housing price along the line," Scientific and Technological Management of Land and Resources, 36(04), 93-103 (2019).

[2] Zhang, L., Liu, J. and Ye, L. Q., "The impact of urban rail transit on housing price along the line based on hedonic model—A case study of Xi'an metro line 2," City, (03), 34-37 (2014).

[3] Wang, X. and Wu, L. X., "The impact of rail transit on urban second-hand housing price—Taking Nanchang metro line 1 as an example," Jiangxi Science, (05), 722-727 (2016).

[4] Ma, W. J., "Analysis on the influence of construction of Zhengzhou metro line 1 on housing price along the line," Modern Commerce and Trade Industry, 40(34), 22-24 (2019).

[5] Liu, Y. S. and Tang, Y., "Analysis on the value-added effect of urban subway on residential price along the line—A case study of Kunming metro line 1 and line 2," Economic Research Guide, (29), 99-101 (2017).

[6] Meng, X. D., Gao, G. Q. and Peng, H. Y., "Research on the impact of subway construction on housing price based on hedonic model," Building Materials Technology and Application, (04), 44-46 (2018).

[7] Yang, Y., "Study on the influencing factors of housing price along Tianjin metro line one based on hedonic model," China Real Estate, (12), 14-22 (2017).

[8] Mohammad, S. I., "The effect of the Dubai Metro on the value of residential and commercial properties," Journal of Transport and Land Use, 10(1), (2017).

[9] Seo, W. and Nam, H. K., "Trade-off relationship between public transportation accessibility and household economy: Analysis of subway access values by housing size," Cities, 87 (2019).

[10] Bae, C.-H. C., Jun, M.-J. and Park, H., "The impact of Seoul's subway line 5 on residential property values," Transport Policy, 10(2), (2003).

[11] Song, D., "An empirical study on the impact of Wuhan metro line 2 on neighborhood housing prices," Modern Business Industry, 37(01), 208-210 (2016).

[12] Ke, Y. Y. and Liu, S. Y., "Spatial and temporal effects of rail transit on housing prices: A case study of Xiamen metro line 1," Construction Economics, 38(09), 90-95 (2017).