Connecting Cities: Assessing the Influence of Metro Line 11 on Housing Prices along the Route between Shanghai and Kunshan

Guanheng Cen, Guodong Zheng, Xiaowen Zhu
Instructor: Sebastiano Manzan

Abstract

The impact of urban rail transit on housing prices has been widely studied, often showing positive effects on property values. This research focuses on Suzhou Metro Line 11, the first metro line to cross the provincial boundary between Shanghai and Jiangsu, and its influence on housing prices in Kunshan. Utilizing OLS model with interactive term included, this study analyzes housing price data from 2022 to 2024, comparing properties within 2 km of the new metro stations to those farther away. The results reveal a localized impact, with a modest 4.9% decrease in housing prices after subway opening observed within 400 meters of the stations, along with other larger changes among longer distance. This result may interpret the construction of Suzhou Metro Line 11 had offset the downward trend of the real estate price in China in recent years. Unlike previous studies, the findings suggest that broader market conditions in post-COVID-19 China may have tempered the expected positive influence of the metro line. The findings offer crucial insights into how newly opened metro lines impact surrounding housing prices in a post-COVID-19 context, providing potential guidance for policymakers, urban planners, and real estate developers in shaping future transportation and housing strategies.

1 Introduction

Urbanization has swept across China since the Reform and Opening-up Era began more than 40 years ago, prompting millions to migrate from rural areas to cities in search of better opportunities. This influx has significantly boosted housing demand, driving rapid increases in housing prices. The Yangtze River Delta (YRD), a key economic hub, has experienced above-average increases in both housing demand and prices. Between 2008 and 2018, housing prices per square meter in YRD cities surged by 146%, compared to the national average increase of 78%, with Shanghai experiencing the highest rise at 203% (Song et al., 2020).

On June 23, 2023, Metro Line 11, connecting Shanghai and Suzhou, commenced operation. This metro line includes 28 newly opened stations, with 26 stations located in Kunshan. This study aims to investigate the impact of the operation of Metro Line 11, the first metro line to cross the provincial boundary between Shanghai and Jiangsu, on second-hand housing prices in Kunshan along its route. We specifically examine how the distance between housing communities and metro stations affects housing prices in these communities.

As the first metro line to serve Kunshan, Metro Line 11 represents a significant development in regional transportation infrastructure. By examining the effects of this intercity metro line, this research sheds light on how such transportation advancements can influence housing markets in suburban, peri-urban, and urban areas in Kunshan. By analyzing housing price trends in Kunshan, this research provides valuable insights for urban planners, policymakers, and real estate developers, aiding them in anticipating the economic and social outcomes of extending metro networks across provincial borders.

To conduct this study, we employ OLS model with interactive term ($POST_{k,i}$) included, despite traditionally using the Hedonic Price Model (HPM). We collected data on housing prices in housing communities along the metro line in Kunshan from 2022 to 2024 to perform a before-and-after analysis of the metro line's impact.

2 Literature Review

2.1 Traditional Hedonic Pricing Model & Difference-in-Difference Method

The impact of newly opened metro stations on housing prices has been the subject of extensive research worldwide. Numerous studies have sought to understand how the introduction of metro infrastructure influences real estate markets, particularly focusing on how distances from the stations affect housing prices. Two methodologies are commonly used in analyzing the increase in housing values: the traditional Hedonic Price Model (HPM) and the Difference-in-Difference (DID) method. The adoption of the HPM (Rosen, 1974) made it simpler to assign property value changes to specific features of properties. The HPM model is widely used in studying the Chinese housing market. For instance, Liu et al. (2015) examined the impact of Nanjing's Metro Lines 1 and 2 on property values. Wen et al. (2018) analyzed how Metro Line 1 in Hangzhou affects housing prices. Xu et al. (2016) investigated property value appreciation near metro rail lines in Wuhan.

However, outside the Chinese housing market, studies often employ the DID method instead of the traditional HPM. For instance, Dubé et al. (2013) utilized a DID estimator to evaluate the impact of a new commuter train service on marginal price changes, accounting for changes in access to stations on Montreal's South Shore in Canada. The reason that studies related to the Chinese housing market frequently use the traditional HPM is the lack of spatio-temporal housing data. The data in the Chinese housing market is mainly cross-sectional data, which suffers from the major problem of endogeneity. The causal relationship between proximity to metro stations and housing prices could be distorted by the endogeneity associated with metro station planning (Tan et al., 2019). Since our study specifically examines how the distance from metro stations affects neighborhood housing prices, we adopt the interactive term to address the endogeneity issue, while collecting spatio-temporal housing data.

Inspired by DID model, we extend the traditional hedonic pricing approach by incorporating the $POST_{k,i}$ (pre- and post-construction) interaction term, which enables us to capture housing price changes across different distance ranges before and after the metro line's construction. By integrating the interaction between POST and Distance (distance groups), our model effectively addresses endogeneity issues, a common challenge in metro station planning, by leveraging the differential price changes over time. Furthermore, by segmenting distance into multiple discrete ranges (e.g., 0-400 meters, 400-800 meters, etc.) and including these interactions, our approach offers a more refined understanding of the heterogeneous impacts of distance on housing prices, providing robust insights into localized effects.

2.2 Scope of Impact

The impact of metro stations on housing prices varies significantly by location. In Atlanta, metro stations influence single-family houses beyond 3 miles (Bowes & Ihlanfeldt, 2001), while in Chicago, the effect on both single-family and multi-family houses is observed within a 1.5-mile radius (McMillen & McDonald, 2004). In contrast, in Daegu, Korea, the impact of new metro stations extends only to 500 meters (Im & Hong, 2018). This variability underscores the need to consider local empirical evidence when determining the scope of impact for metro stations. In China, the scope of impact typically ranges from within 1 km to a maximum of 2 km. For example, Xu et al. (2016) reported significant price increases of 16.7% within 100 meters and 8% between 100 and 400 meters from stations in Wuhan. In Nanjing, Liu et al. (2015) found significant appreciation within 1.5 km of the metro lines, with no notable effects beyond this distance. Sun et al. (2016) observed significant appreciation within 1 km of Line 3 in Tianjin, and Dai et al. (2016) found that in Beijing, transfer stations impacted housing prices up to 1200-1400 meters, while non-transfer stations affected prices up to 1 km. Consequently, our study on Metro Line 11 in Kunshan, a non-transfer line, extends the study area to housing communities up to 2 km (straight-line distance) from the metro stations.

3 Data and Model

3.1 Study Area and Data Source

This study focuses on how the opening of Suzhou Metro Line 11 in June, 2023 affects neighborhood housing prices in Kunshan. To approach this question, We classify housing communities into a treatment group (within straight-line distance of 2 km from the subway station) and a control group (beyond 2 km), and then further divide the treatment group into five subgroups based on each community's straight-line distance to its nearest station. We obtain transaction data for these communities from the Chengshu system (wanshudata.com) between the third quarter of 2022 and the second quarter of 2024, a timeframe that includes both pre- and post-opening periods of the metro line to enable before-and-after analysis. In total, we collect 11,154 transaction records, providing a rich dataset to examine how proximity to Metro Line 11 affects local housing prices.

3.2 Variable Selection

Variable selection is critical for capturing the spatial and temporal dynamics of housing prices influenced by metro line construction. In this model, we divide the variables into three main categories: **Distance Dummies**, **Control Variables**, and **Interaction Terms**.

1. Distance Dummies (Distance$_{k,i}$)

The distance dummies reflect the proximity of properties to the metro station. Five dummy variables are created to represent specific distance bands:

- $0 - 400$ meters,
- $400 - 800$ meters,
- $800 - 1200$ meters,
- $1200 - 1600$ meters,
- $1600 - 2000$ meters.

Properties located more than 2000 meters away before the opening of subway line serve as the **base group** for comparison. These variables measure the spatial variation in housing prices **before the metro opening** and establish the baseline spatial patterns.

2. Control Variables

Control variables account for other intrinsic property characteristics that influence housing prices, ensuring the metro's effects are isolated. The included control variables are:

- **House Area** ($Area_i$): The size of the property in square meters.
- **Quantity of Bedrooms** ($QBedroom_i$): The number of bedrooms in the property.

- **Floor** ($Floor_i$): A dummy variable, 1 indicates a high-rise property and 0 otherwise.
- **Decoration** ($Decoration_i$): A dummy variable, where 1 indicates that the property is well-decorated and 0 otherwise.
- **Orientation** ($Orientation_i$): A dummy variable, where 1 indicates that the property is north-facing and 0 otherwise.
- **Age** (Age_i): The number of years from the property's construction to its sale date.

These control variables help mitigate omitted variable bias and ensure that the effects of metro accessibility are not confounded by unobserved housing characteristics.

3. Interaction Terms ($POST_{k,i}$)

Interaction terms are constructed to evaluate the post-construction impact of the metro line on housing prices across distance bands. Specifically:

$$\begin{aligned}
POST_0\text{-}400, i &= POST_i \cdot \text{Distance}_{\text{0-400},i}, \\
POST_400\text{-}800, i &= POST_i \cdot \text{Distance}_{\text{400-800},i}, \\
POST_800\text{-}1200, i &= POST_i \cdot \text{Distance}_{\text{800-1200},i}, \\
POST_1200\text{-}1600, i &= POST_i \cdot \text{Distance}_{\text{1200-1600},i}, \\
POST_1600\text{-}2000, i &= POST_i \cdot \text{Distance}_{\text{1600-2000},i}.
\end{aligned}$$

Here, $POST_i$ is a dummy variable equal to 1 if the transaction occurred **after the metro construction** and 0 otherwise. These interaction terms allow us to measure the incremental effects of metro construction on housing prices in each distance band relative to the base group and pre-construction period.

3.3 Empirical Model

The empirical model is specified as follows:

$$\begin{aligned}
\ln P_i = \beta_0 &+ \beta_1 \text{Distance}_{0-400,i} + \beta_2 \text{Distance}_{400-800,i} + \beta_3 \text{Distance}_{800-1200,i} \\
&+ \beta_4 \text{Distance}_{1200-1600,i} + \beta_5 \text{Distance}_{1600-2000,i} \\
&+ \beta_6 QBedroom_i + \beta_7 Area_i + \beta_8 Floor_i + \beta_9 Orientation_i \\
&+ \beta_{10} Age_i + \beta_{11} Decoration_i \\
&+ \delta_1 POST_0\text{-}400, i + \delta_2 POST_400\text{-}800, i + \delta_3 POST_800\text{-}1200, i \\
&+ \delta_4 POST_1200\text{-}1600, i + \delta_5 POST_1600\text{-}2000, i + \epsilon_i
\end{aligned}$$

Where:

- $\ln P_i$: Natural logarithm of P_i, transaction price per square meter for property i, stabilizing variance and allowing coefficients to be interpreted as percentage changes,

- $\text{Distance}_{k,i}$: Dummy variables representing distance bands ($k = 0\text{-}400, 400\text{-}800, 800\text{-}1200, 1200\text{-}1600, 1600\text{-}2000$) for property i, where the base group consists of properties located more than 2000 meters away.

- $POST_{k,i}$: Interaction terms ($POST_i \cdot \text{Distance}_{k,i}$), capturing the additional price effects after metro construction for each distance band,

- $QBedroom_i, Area_i, Floor_i, Orientation_i, Age_i, Decoration_i$: Control variables for property characteristics,

- ϵ_i: Error term capturing unobserved factors.

4 Empirical Application

4.1 Rationale and Applicability of the Model

The proposed model is built to comprehensively evaluate the causal impact of metro line construction on housing prices, incorporating both **spatial variation** and **temporal dynamics** while addressing potential confounding factors. The structure of the model ensures that the analysis is robust and logically sound for identifying the effects of metro accessibility. Below, we highlight the specific advantages of the model through three key aspects: **spatial baseline effects**, **causal inference through interaction terms**, and **control for confounding factors**.

4.1.1 Capturing Baseline Spatial Effects through Distance Dummies

The inclusion of distance dummies $\text{Distance}_{k,i}$, where $k = 0\text{-}400, 400\text{-}800, \ldots, 1600\text{-}2000$, allows us to measure the **pre-construction differences** in housing prices across different proximity bands. This spatial breakdown achieves two main objectives:

- **Baseline Comparisons**: By dividing properties into distinct bands and comparing them to the **base group** (properties located more than 2000 meters away before the subway opening), we account for pre-existing price variations that may arise from different locations, accessibility to existing infrastructure.

- **Identification of Spatial Heterogeneity**: The coefficients β_k (where $k = 1, \ldots, 5$) associated with each distance band capture the spatial heterogeneity in housing prices before the metro became operational. This step ensures that any observed post-construction changes can be interpreted relative to a clearly defined spatial baseline.

The clear separation of properties into spatial bands provides a solid foundation for interpreting both pre-construction price patterns and the changes induced by the metro's opening further.

4.1.2 Measuring Causal Effects through Interaction Terms

To identify the **causal impact** of metro construction, the model introduces interaction terms $POST_{k,i} = POST_i \cdot \text{Distance}_{k,i}$, where $POST_i$ is a dummy variable equal to 1 if the transaction occurred after the metro line opened and 0 otherwise. These terms serve two critical purposes:

- **Spatially-Varying Effects**: By interacting the post-construction indicator $POST_i$ with each distance dummy, the model allows the effects of the metro to **vary spatially** across different proximity bands. This flexibility ensures that we capture localized price responses, which may differ depending on how close properties are to the station.

- **Causal Interpretation and trend of δ_k**: The coefficients δ_k (where $k = 1, \ldots, 5$) associated with the interaction terms represent the **incremental impact** of the metro's opening on housing prices. A positive δ_k indicates a price premium for properties that not only offsets any potential negative impact from external market conditions but also exceeds it, while a negative δ_k may reflect the non-negligible negative impact of external market conditions outweighing the positive accessibility benefits of new metro. Also, the trend of coefficient δ_k may provide insights into how the metro's impact changes across distance bands.

By introducing these interaction terms, the model mimics a **Difference-in-Differences (DID)-like framework**, where price changes before and after the metro's construction are compared across treatment group (closer properties) and control group (faraway properties). This setup enhances the model's ability to identify the **causal effects** of the metro while accounting for spatial variations.

4.1.3 Addressing Confounding Factors through Control Variables

To ensure that the metro's impact on housing prices is not confounded by other property-level characteristics, the model incorporates a comprehensive set of control variables:

$$QBedroom_i,\ Area_i,\ Floor_i,\ Orientation_i,\ Age_i,\ Decoration_i$$

These variables account for intrinsic property features that naturally influence housing prices:

- **Size and Functional Space**: House area and quantity of bedrooms reflect the utility and desirability of the property.
- **Property Quality**: Decoration status and floor level account for quality differences that can affect prices.
- **Depreciation Effects**: Property age captures the impact of depreciation on older houses.
- **Orientation Preferences**: North-facing properties (captured through the orientation dummy) may command higher or lower prices depending on buyer preferences.

By including these controls, the model ensures that any observed changes in housing prices across spatial bands are driven primarily by metro accessibility and not by unobserved property-level heterogeneity.

4.2 Interpretation of Results

We use the method of Ordinary Least Squares (OLS) to run the regression model. Given the potential for heteroskedasticity caused by variations in house-specific characteristics and distance bands, we adjust the standard errors using the HC1 robust covariance estimator. This ensures that our results are reliable and heteroskedasticity-robust. From the results shown in the table below, the relationship between house prices and proximity to the nearest subway station, post-renovation effects, as well as other house-specific characteristics, can be seen. Specifically, houses within 0–400 meters from the nearest subway station show that with a coefficient of 0.1150, house prices would be 11.50% higher than elsewhere in the base category of over 400 meters from the nearest subway station before construction. For houses within 400–800 meters, the coefficient of 0.1144 infers an increase in housing prices by 11.44%. When the distance falls within 800–1200 meters, the housing price is expected to increase by 9.62%, as suggested by the coefficient 0.0962. For houses between 1200–1600 meters, the coefficient of 0.0755 predicts a rise of 7.55% in housing prices.

Table 1: OLS Regression Results

Model Summary			
Dep. Variable:	log_price	**Log-Likelihood:**	-2628.1
R-squared:	0.151	**No. Observations:**	11154
Adj. R-squared:	0.149	**AIC:**	5290
Method:	Least Squares	**BIC:**	5415
F-statistic:	111.2	**Df Residuals:**	11137
Date:	Sat, 14 Dec 2024	**Df Model:**	16
Prob (F-statistic):	0.00	**Covariance Type:**	HC1

Regression Coefficients

	coef	**std err**	**z**	**P> $\lvert z\rvert$**	**[0.025**	**0.975]**
const	9.6380	0.019	509.403	0.000	9.601	9.675
Distance$_{0-400}$	0.1150	0.023	5.002	0.000	0.070	0.160
Distance$_{400-800}$	0.1144	0.012	9.487	0.000	0.091	0.138
Distance$_{800-1200}$	0.0920	0.013	7.295	0.000	0.070	0.122
Distance$_{1200-1600}$	0.0755	0.012	6.158	0.000	0.051	0.100
Distance$_{1600-2000}$	0.0680	0.014	4.469	0.000	0.029	0.018
Area	0.0025	0.003	0.910	0.363	-0.000	0.001
Qbedroom	0.0782	0.007	10.549	0.000	0.062	0.090
Floor	-0.0446	0.006	-7.719	0.000	-0.057	-0.032
Decoration	0.0590	0.007	8.792	0.000	0.041	0.065
Orientation	-0.0353	0.007	-5.180	0.000	-0.049	-0.022
Age	-0.0060	0.001	-8.674	0.000	-0.007	-0.005
POST_0-400	-0.0490	0.034	1.438	0.150	-0.116	0.018
POST_400-800	-0.2095	0.014	-14.910	0.000	-0.237	-0.182
POST_800-1200	-0.2602	0.016	-16.413	0.000	-0.291	-0.229
POST_1200-1600	-0.2393	0.011	-20.932	0.000	-0.262	-0.217
POST_1600-2000	-0.1575	0.010	-15.222	0.000	-0.178	-0.137

The coefficients for the post-renovation variables ($POST_{k,i}$) indicate a huge decrease in the prices of housing. Coefficients of -0.0490 for houses within 0–400 meters of the subway station show that there is a fall in price by 4.90% after the construction of the subway, but it has been insignificant statistically at the 5% confidence level. In contrast, for houses located 400–800 meters and 800–1200 meters post-renovation, housing prices show significant declines of 20.95% (coefficient: -0.2095) and 26.20% (coefficient: -0.2620), respectively. Similarly, houses within 1200–1600 meters post-renovation experience a 23.93% decrease (coefficient: -0.2393), while those within 1600–2000 meters exhibit a 15.75% decrease (coefficient: -0.1575) compared to the base category. From the result shown above, we can tell that the decrease in the price of the house is becoming larger as the house moves further to the nearest subway station, which shows that the construction of the subway has offset the decrease of China's negative house market at some level. Overall, the closer it is to the metro, the better the offset is, indicating the positive effect of metro line 11 on the housing price.

4.3 Limitation of the Model

Our project primarily focuses on evaluating the effect of the opening of a single subway line, specifically Kunshan's Metro Line 11. We chose this line because it is the only subway line in Kunshan, thereby eliminating interference from other subway lines. However, our model does have some limitations. For instance, factors such as proximity to the central business district (CBD), hospitals, schools, and other urban amenities were not included due to data constraints. These omitted variables likely influence housing prices, potentially introducing omitted variable bias in our results. Additionally, serial correlation may exist, as current housing prices can influence future housing prices, which is another aspect not fully accounted for in our analysis.

5 Conclusion

These findings run counter to several previous studies that report significant general impacts of new metro lines on surrounding housing prices over a wide area of influence. Unlike earlier studies that often report broader and more sustained increases in property values due to improved metro accessibility, the impact in Kunshan appears localized and limited. The reasons could be attributed to broader economic conditions in China, particularly the impact of COVID-19. The pandemic has brought significant uncertainty and economic fluctuation, disrupting housing market dynamics and dampening the expected impacts of the metro line's opening.

It is plausible that these broader economic challenges diminished the benefits of the improved transportation network. This highlights the need for additional research on how large economic shocks, such as global pandemics, alter the conventional relationship between transportation infrastructure and housing markets. Future studies can confirm whether Kunshan's case is an exception or part of a broader trend across China in the post-pandemic era.

In summary, Metro Line 11 did not lead to a significant increase in housing prices near the stations as might have been expected. The results suggest that the advantages of metro accessibility were offset by other factors, such as broader economic uncertainties and the impact of the COVID-19 pandemic. This indicates that macroeconomic conditions and external economic shocks play a dominant role in shaping housing price dynamics, potentially dampening the expected benefits of improved metro accessibility.

References

[1] Bowes, D. R., & Ihlanfeldt, K. R. (2001). Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics*, 50(1), 1–25. https://doi.org/10.1006/juec.2001.2214

[2] Dai, X., Bai, X., & Xu, M. (2016). The influence of Beijing rail transfer stations on surrounding housing prices. *Habitat International*, 55, 79–88. https://doi.org/10.1016/j.habitatint.2016.02.008

[3] Dubé, J., Thériault, M., & Des Rosiers, F. (2013). Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice*, 54, 49–66. https://doi.org/10.1016/j.tra.2013.07.015

[4] Im, J., & Hong, S. H. (2017). Impact of a new subway line on housing values in Daegu, Korea: Distance from existing lines. *Urban Studies*, 55(15), 3318–3335. https://doi.org/10.1177/0042098017741412

[5] Liu, K., Wu, Q., & Wang, P. (2015). Quantitative analysis of the impact of urban rail transit on housing prices: A case study of Nanjing Metro Line 1 and Line 2. *Resources Science*, 37(1), 133–141.

[6] McMillen, D. P., & McDonald, J. (2004). Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983–1999. *Real Estate Economics*, 32(3), 463–486. https://doi.org/10.1111/j.1080-8620.2004.00099.x

[7] Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34–55.

[8] Song, W., Chen, Y., Sun, J., & He, M. (2020). Spatial differentiation of urban housing prices in the integrated region of the Yangtze River Delta. *Acta Geographica Sinica*, 75(10), 2109–2125. https://doi.org/10.11821/dlxb202010006

[9] Sun, H., Wang, Y., & Li, Q. (2016). The Impact of Subway Lines on Residential Property Values in Tianjin: An Empirical Study Based on Hedonic Pricing Model. *Discrete Dynamics in Nature and Society*, 2016, 1–10. https://doi.org/10.1155/2016/1478413

[10] Tan, R., He, Q., Zhou, K., & Xie, P. (2019). The effect of new metro stations on local land use and housing prices: The case of Wuhan, China. *Journal of Transport Geography*, 79, 102488. https://doi.org/10.1016/j.jtrangeo.2019.102488

[11] Wen, H., Gui, Z., Tian, C., Xiao, Y., & Fang, L. (2018). Subway Opening, Traffic Accessibility, and Housing Prices: A Quantile Hedonic Analysis in Hangzhou, China. *Sustainability*, 10(7), 2254. https://doi.org/10.3390/su10072254

[12] Xu, T., Zhang, M., & Aditjandra, P. T. (2016). The impact of urban rail transit on commercial property value: New evidence from Wuhan, China. *Transportation Research Part A: Policy and Practice*, 91, 223–235. https://doi.org/10.1016/j.tra.2016.06.026