Effect of Proximity to Metro Stations and Distance from City Center on Rental Prices in Delhi-NCR

Ayaan Dutt

Group Members: Shreya Anand, Raheem Manoj, Bhavana Gudnavar

Professor: Manvi Sharma
TFs: Arastu Pandey, Jishnu Borgohain,
Nupur Agrawal, Sarthak Mohanty

Date of Submission: May 11th, 2024

Introduction

Urban rental markets are influenced by a complex interaction of geographic, economic, and infrastructural factors. In a rapidly expanding metropolis like Delhi-NCR, it is crucial to understand the determinants for housing prices for effective policy-making, urban planning, and investment strategy. The accessibility of public transport and proximity to economic hubs are likely to play a pivotal role in influencing residential rental prices. This study examines the extent to which (1) distance from the central business district (CBD) and (2) proximity to the nearest metro station affect residential rental prices per square foot in the Delhi-NCR region.

The accessibility of public transport, especially a well-connected transit system such as the Delhi metro, often enhances mobility and reduces commuting costs, potentially increasing property values with greater connectivity. This is especially relevant for residential rental pricing since tenants are usually of working age and are attracted by better access to economic and commercial hubs. For similar reasons, the proximity to the CBD may increases the desirability and utility of a property, thereby increasing its market value.

Empirical studies from various global cities have consistently shown that properties located closer to public transit and economic hubs tend to command higher rental and sale values. However, while such trends are well-documented in cities like Shanghai, London, or New York, relatively fewer studies have rigorously quantified these effects in Indian cities. With the Delhi Metro emerging as one of the largest and most efficient rapid transit systems in the world, its influence on real estate markets is both timely and significant. As the region continues to expand outward, it is crucial to understand whether rental premiums associated with metro access vary depending on a property's location relative to the city centre as well.

Literature Review

The CBD typically represents the core of economic and administrative activity in a city. According to classical urban economic theory — most notably the monocentric city model (Alonso, 1964; Muth, 1969) — property values tend to decrease with increasing distance from the CBD due to rising commuting costs and reduced access to jobs, services, and amenities. Empirical studies reinforce this pattern. For instance, Chen & Hao (2008) found a strong negative correlation between distance from the CBD and housing prices in Shanghai, China, using a hedonic pricing model.

This pattern is likely to hold in Delhi-NCR, where central areas such as Connaught Place, Barakhamba Road, and the surrounding commercial zones form the economic nucleus. Given Delhi-NCR's polycentric growth, the strength of the CBD effect may vary across sub-regions, but the general principle remains: greater accessibility to the core urban economy typically imposes a premium on rental pricing.

Access to public transportation, particularly rapid transit systems such as metros have also been shown to positively influence real estate prices. Several studies in metropolitan cities across the globe including Washington DC (Grass, 1992), London (Gibbons & Machin, 2004) and Beijing (Zhou et. al., 2022) have documented that properties located closer to metro stations exhibit higher sale and rental prices, even after controlling for other factors. Similar analyses have been carried out by Singhal & Tyagi (2021) in Delhi-NCR in which they used hedonic price models to control for other characteristics, such as structure, environment, location and neighbourhood to show that the metro induced a price increase of Rs.246-732 (station-wise) in the vicinity of 500 meters of the station.

Some studies have also examined interaction effects - how the benefit of metro proximity may vary depending on a property's distance from the city centre. For example, research suggests that the metro transit premium may be more pronounced in suburban or peripheral areas, where alternative transportation options are limited and commuting distances are long (Dai et. al, 2016). In contrast, in well-connected central districts, the marginal value of being near a metro station may be lower due to overlapping transport options and better infrastructure. Given Delhi-NCR's sprawling geography and the decentralized distribution of metro lines, investigating such interaction effects may reveal that distant suburbs like Noida Extension or Gurugram derive greater relative value from metro access than those located near Connaught Place or South Delhi.

Study Design

Based on past literature, we make the following hypotheses:

- **Hypothesis 1**: There exists a negative correlation between distance to the nearest metro station for a residential property and the rent paid per square foot. We expect this correlation to eventually drop off beyond a certain distance, implying a non-linear relationship between the variables.
- **Hypothesis 2**: There exists a negative correlation between the distance of a property from the Central Business District and the residential rent paid per square foot. Here too, we expect the rent to eventually asymptote as you approach the outskirts of the city, suggesting a non-linear relationship.

The response variable has been quantified in terms of the rental price per square foot (R), based on government records, market listings and survey data. We restrict ourselves to residential properties to avoid confounding factors that may arise due to alternative priorities of non-residential property tenants and owners. The distance between properties and their nearest metro station (D_m) can be measured (in meters) along the shortest commute distance by road, using remote sensing and Google Maps. Finally, using Connaught Place as the CBD location for Delhi-NCR, it is possible to measure the straight-line distance (in kilometres) between properties and the centre of the Connaught Place circle using remote sensing to determine the distance to the CBD or city centre (D_{CBD}).

We expect the distribution of rental prices for a given neighbourhood to be right-skewed since property values reflect the income distribution of the residents. Studies show that housing prices serve as an effective proxy for income distribution which has a right-skew power-law tail which decays to a finite upper limit (Liu, et. al., 2024). This is likely due to the fact that any given area is largely composed of more low-to-mid income residents and a few very high-income residents who can afford substantially higher rents. However, this can be dealt with by log-transforming the rental data to normalize it. We can also sample the data such that the skew-ness is less pronounced, allowing us to use a simple LM to analyse it.

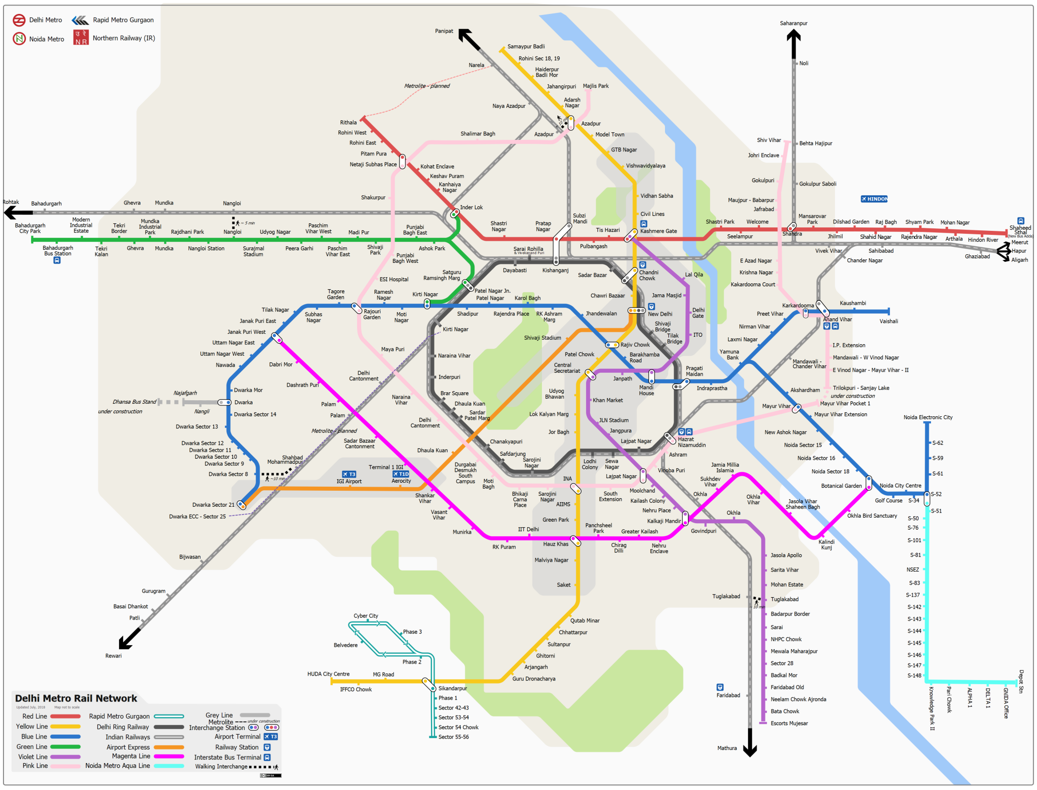

Figure 1: The Delhi Metro shows a greater density of metro stations closer to the CBD (Rajiv Chowk)
Source: DMRC

We may also expect there to be a correlation between D_m and D_{CBD} due to variable density of metro stations across Delhi-NCR. Since the density of metro stations for the Delhi metro increases closer towards the city centre (fig 1), we would expect the average distance of a property to a metro station to reduce. This can be resolved by sampling data uniformly both close to the CBD and in the suburbs and outskirts, thereby removing any bias in the data and decoupling the variables.

Model Specification

Based on the study design and objectives, we propose a linear regression model with a log-transformed rent to describe our system. For simplicity, we have ignored the interaction effects from $D_m \times D_{CBD}$. The following equation describes the predicted rent for a given property:

$$\ln(R) = \beta_0 + \beta_1 D_m + \beta_2 D_{CBD} + \epsilon \tag{1}$$

Where,

- R is the residential rental price (in rupees per square foot).
- D_m is the distance from the property to its nearest metro station (in meters).
- D_{CBD} is the distance from the property to the CBD (in kilometres).
- β_0, β_1 and β_2 are the intercept and regression coefficients. Our hypotheses predict a negative correlation for both independent variables, which would imply that $\beta_1, \beta_2 < 0$.
- ϵ is the error due to fluctuations that may arise from other factors. We assume that this error is stochastic and therefore has a mean of 0 and constant variance.

This model provides a framework to measure the direct effects of the distance to the CBD and proximity to metro stations on rental prices in Delhi-NCR. The sign of the coefficients β_1 and β_2 will be used to test our hypotheses and their magnitude will determine the overall effect size.

Analysis

We first tested the distribution of the rental data and contrary to our expectations, we found that it was normally distributed instead right-skewed (fig 2, left). This could be achieved through clever sampling and is not necessarily representative of the natural distribution of rent. We allow for this since the objective of this study is to determine causation rather than observe patterns. Running a Shapiro–Wilk test for normality, we got a p-value of 0.287, which means that we cannot reject the null hypothesis that the data is normally distributed (i.e. there is no significant deviation from normality). We can also see that the normal Q-Q plot returns a reasonable fit (fig 2, right). Hence, we did not log-transform the rent and can run a simple linear regression model.

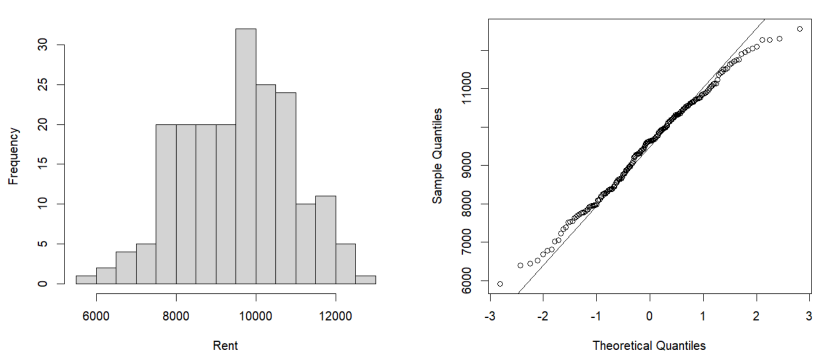

Figure 2: The rental price data is normally distributed
as seen by the histogram (left) and normal Q-Q plot (right)

Running an ordinary least squares (OLS) LM method on R using rent as the response variable and distance to CBD and metro proximity as the predictor variables, we received the following output:

```
Call:
lm(formula = rent ~ m.dist + cc.dist, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-1124.84  -366.93    -6.58   302.57  1557.00

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.208e+04  8.178e+01  147.67   <2e-16 ***
m.dist      -1.232e+00  4.197e-02  -29.35   <2e-16 ***
cc.dist     -9.320e+01  3.999e+00  -23.30   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 495 on 197 degrees of freedom
Multiple R-squared:  0.8706, Adjusted R-squared:  0.8693
F-statistic: 662.9 on 2 and 197 DF,  p-value: < 2.2e-16
```

Based on the coefficients, we can find the predicted rent from the expression:

$$R = 12080 - 1.232D_m - 93.2D_{CBD} \pm 495 \tag{2}$$

The multiple R-squared value of 0.8706 means that the model explains 87.06% of the variance in rental prices, which is a very good fit. The adjusted R-squared value is slightly lower at 86.93%, which accounts for the number of predictors in the model. The results are highly significant, with p-values well below 0.05 ($p < 2 * 10^{-16}$). The regression coefficients can be interpreted in the following manner:

- $\beta_0 = 12,080 \pm 82$ (in rupees per square foot) is the intercept, which is equivalent to the rent in the absence of any effects of metro proximity and CBD distance. Though this is an unrealistic scenario, it provides us with a reference point for the model.

- $\beta_1 = -1.232 \pm 0.042$ (in rupees per metre per square foot) is the coefficient for metro distance. This means that for every additional meter away from the nearest metro station, the rental price decreases by $Rs.1.23$ per square foot, holding CBD distance constant (fig 3, left). This confirms a negative correlation between metro distance and rent.

- $\beta_2 = -93.2 \pm 4.0$ (in rupees per kilometre per square foot) is the coefficient for distance from CBD. Hence, for every additional kilometre away from the CBD, the rental price decreases by $Rs.93.20$ per square foot, holding metro distance constant (fig 3, right). This confirms a negative correlation between CBD distance and rent.

- $\epsilon = 495$ (in rupees per square foot) is the residual standard error, which is the average size of the error in the prediction. We attribute this to random factors that cause local fluctuations in rental prices.

Using ggplot2 to plot the regression lines with error, we produced the following plots:

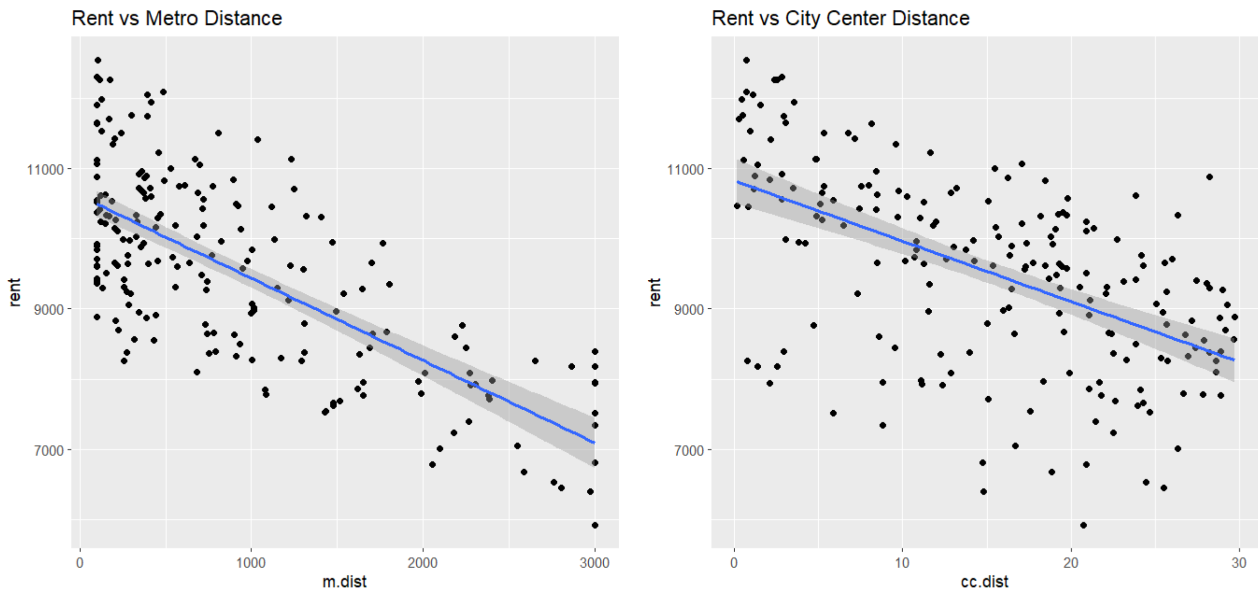

Figure 3: Rent shows a negative correlation with both distance from CBD and distance to nearest metro station

Running a full model check we got the following outputs:

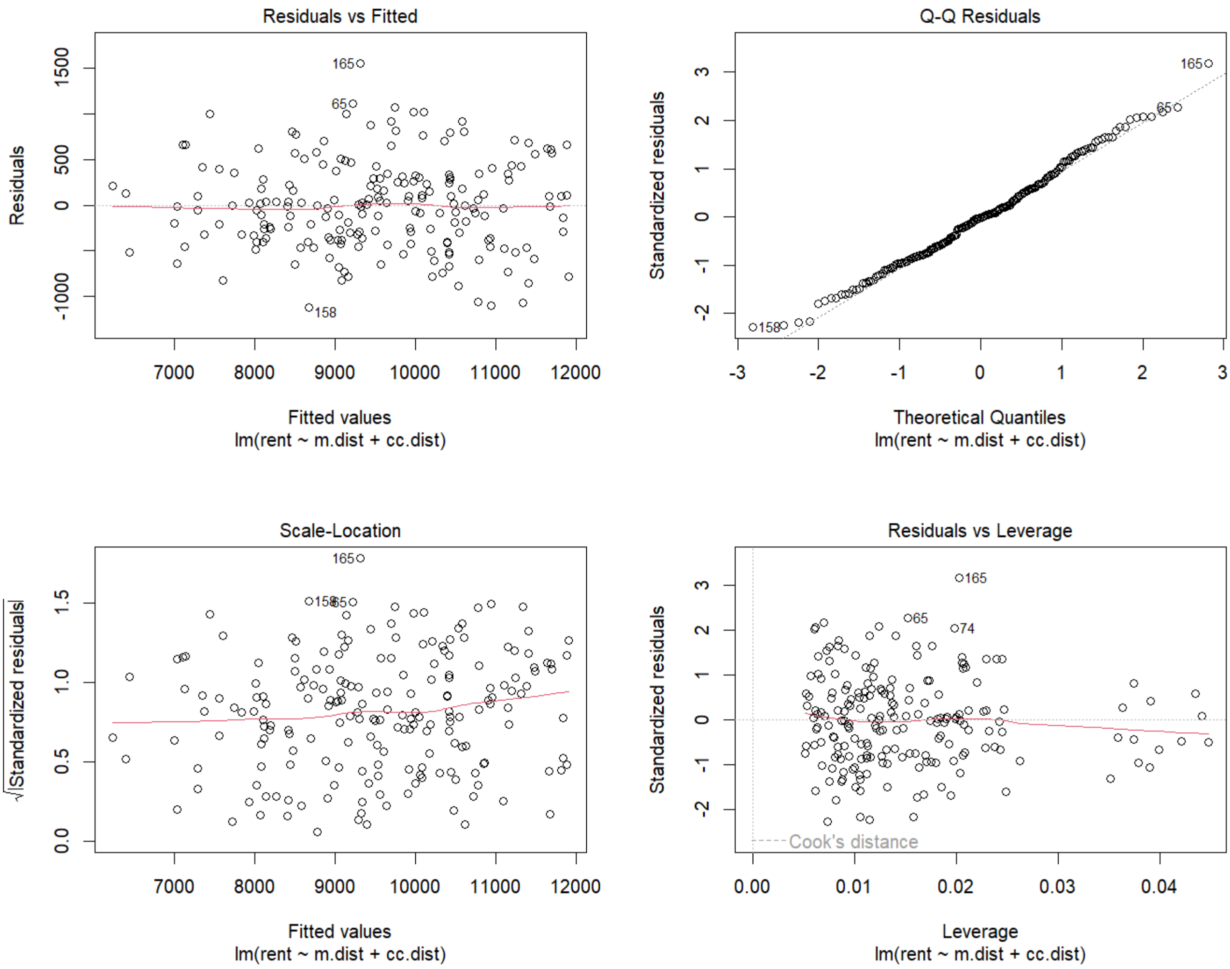

Figure 4: Model checking plots for $R = \beta_0 + \beta_1 D_m + \beta_2 D_{CBD} + \epsilon$

The residual analysis shows that the data is fairly centred with the median value of -6.58. The range from -1124 to +1557 indicates the spread of prediction errors, with a few extreme cases. The residuals vs fitted plot shows a uniform distribution across values which supports linearity and constant variance of the data. The Q-Q residual plot also shows a good fit along the 45-degree line, confirming the normality of the residuals. The scale-location plot checks for homoscedasticity (similar to plot 1) and the flat horizontal line suggests a constant variance. Finally, the residuals vs leverage plot can be used to identify excessively influential outliers, which in our case are negligible. Hence, the model seems to be an appropriate choice for the given data.

Discussion

These findings are consistent with the monocentric city model and empirical results from other cities, but their quantification in the specific context of Delhi-NCR adds localized policy relevance. The adjusted R-squared value of 86.9% indicates that a large proportion of the variability in rental prices can be explained by these two geographic variables alone, highlighting their importance. Despite its strengths, this study has some limitations:

- The exclusion of interaction terms means the analysis does not account for varying effects of metro proximity across different distances from the CBD.
- The use of straight-line distance for CBD distance rather than actual travel times may oversimplify real-world accessibility.
- Other potentially relevant predictors such as neighbourhood quality, amenities, building age, and income levels were not included due to data limitations.
- The sampling approach used to normalize the data may limit external validity, as real-world rental distributions are typically right-skewed.

One interesting observation is that the data, contrary to initial expectations, was approximately normally distributed, likely due to sampling methods that controlled for extreme values. While this limits generalizability to the broader market, it improves the clarity of causal inference.

Conclusion

This study confirms that public transport and commercial hub accessibility, in particular, proximity to metro stations and distance to the central business district, plays a significant role in determining residential rental prices in Delhi-NCR. Both variables were found to be negatively correlated with rent, with statistically significant coefficients and a strong model fit. These results offer insight into the primary determinants of residential rental prices and the rate and which they vary.

Future research could incorporate interaction terms, more detailed spatial and socioeconomic variables, and alternative models to deepen understanding of these complex relationships. Nonetheless, this study offers a robust, interpretable model with clear policy and investment relevance for the Delhi-NCR housing market.

References

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent.* Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674730854

Chen, J., & Hao, Q. (2008). *The impacts of distance to CBD on housing prices in Shanghai: a hedonic analysis.* Journal of Chinese Economic and Business Studies, 6(3), 291–302. https://doi.org/10.1080/14765280802283584

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

Gibbons, S., & Machin, S. (2004). *Valuing rail access using transport innovations.* Journal of Urban Economics, 57(1), 148–169. https://doi.org/10.1016/j.jue.2004.10.002

Grass, R. G. (1992). *The estimation of residential property values around transit station sites in Washington, D.C.* Journal of Economics and Finance, 16(2), 139–146. https://doi.org/10.1007/bf02920114

Liu, J., Farahani, H., & Serota, R. A. (2024). *Exploring distributions of house prices and house price indices.* Economies, 12(2), 47. https://doi.org/10.3390/economies12020047

Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use.* Chicago and London: The University of Chicago Press.

Singhal, S., & Tyagi, Y. (2021). *Analyzing the influence of metro stations on commercial property values in Delhi: A hedonic approach.* Real Estate Management and Valuation, 29(4), 10–22. https://doi.org/10.2478/remav-2021-0026

Zhou, Y., Tian, Y., Jim, C. Y., Liu, X., Luan, J., & Yan, M. (2022). *Effects of public transport accessibility and property attributes on housing prices in polycentric Beijing.* Sustainability, 14(22), 14743. https://doi.org/10.3390/su142214743