JICA Research Institute

JICA-RI Working Paper

Study on the Methodology of Impact Analysis for Infrastructure Projects

# Heterogeneous Effects of Urban Public Transportation on Employment by Gender: Evidence from the Delhi Metro

**Mai Seki and Eiji Yamada**

No. 207

March 2020

JICA Research Institute

JICA Research Institute

10-5 Ichigaya Honmura-cho

Shinjuku-ku

Tokyo 162-8433 JAPAN

TEL: +81-3-3269-3374

FAX: +81-3-3269-2054

# Heterogeneous Effects of Urban Public Transportation on Employment by Gender: Evidence from the Delhi Metro*

Mai Seki† and Eiji Yamada‡

**Abstract**
The Delhi Metro is one of the leading examples of a recent urban mass transit infrastructure project in a developing country where women have traditionally suffered from constrained mobility. In this paper, we analyze the effects of the Delhi Metro on the work participation rate of women and men, using a three-period (1991, 2001, and 2011) panel data of township-level zones within the city of Delhi. While the data has limitations in understanding the characteristics of individual residents in detail, we employ a difference-in-differences estimation controlling for a location fixed-effect, with a parallel trend test. The results suggest that the proximity to the Delhi Metro stations significantly increases the female work participation rate (WPR), whereas its effect on the male WPR is ambiguous with the potential to have an opposite sign. While there are number of potential mechanisms that can deliver this result, we develop a theoretical urban commuting model and argue that a larger reduction in the commuting cost for females (by offering a safer commuting mode of transportation, for example) can generate the quantified patterns of the effects on the WPR. Overall, our results relate to the literature on the quantification of the contribution of urban transport infrastructure towards inclusive growth and poverty reduction.

**Keywords:** India, gender gap, equilibrium commuting model

**JEL**: O18, J20, R00

*This paper was prepared under a JICA-RI project titled "Study on the Methodology of Impact Analysis for Infrastructure Projects". We are thankful to the seminar participants at Lisbon 2017, GRIPS 2017, Niigata 2017, ISEC Bengaluru 2018, JEA Kobe 2019, and JASID Tokyo 2019. We are also grateful to Prof. Takashi Kurosaki, Prof. Kensuke Teshima, and two anonymous reviewers for their valuable comments for the earlier versions of this paper. The study was prepared by the authors in their own personal capacity. The opinions expressed in this article are the authors' own and do not reflect the official positions of either the JICA Research Institute or JICA. All errors are ours.

†Ritsumeikan University (maiseki@fc.ritsumei.ac.jp)

‡JICA Research Institute (Yamada.Eiji@jica.go.jp)

# 1. Introduction

In the past seventy years, the share of urban dwellers has steadily increased in developing countries, and this trend will continue in the coming decades (United Natations 2019). India is one of the main contributors to the global urban population growth, which is projected to add 416 million urban residents by 2050. To mitigate traffic congestion accompanied by the continuing urbanization, many countries including India are investing in urban public transportation systems. While the overall mobility of residents improves and city production capacities expand, gender inequality of mobility in urban areas remains an unresolved issue (Peters 2013; Uteng 2011; Hyodo et al. 2005). According to previous studies, women in the urban areas of the developing countries go out of home less frequently, and depend more on public transportation than men. The provision of safe and accessible public transportation could potentially improve female mobility, a necessary conditoin for their further active participation in the economy.

In fact, a gender mainstreaming in the infrastructure projects of developing countries has gained attention from policy makers over the past decade (Asian Development Bank 2013; African Development Bank Group 2009; UN Women 2014; World Bank 2010). However, there is still only a limited amount of research quantifying the development impact, especially on how women and men are differentially affected by urban transport development. [1] There are studies that have discussed gender heterogeneity in commuting time to work and its impact on labor supply (Gutiérrez-i-Puigarnau and Ommeren 2010; Gimeneznadal and Molina 2014; Gimenez-Nadal and Molina 2016; Zax 1991; Black, Kolesnikova, and Taylor 2014); however, they do not necessarily focus on public transportations in a

1. There is a large literature on the effect of subways on employment density in the developed countries such as that by Redding and Turner (2015). However, very few impact evaluations of urban transportation exist in the developing countries. Majority features rural roads and some major studies discuss inter-city highways or railroads (Seki, 2016).

given country context, except the ones by Kawabata and Abe (2018) and Gaduh, Gracner, and Rothenberg (2018). Kawabata and Abe (2018) analyze the commuting and labor supply patterns of married couples, resident in the greater Tokyo metropolitan area using GIS. Gaduh, Gracner, and Rothenberg (2018) estimate an equilibrium model of commuting choices with endogenous commuting time to assess the impact of counterfactual transportation policies, using the data collected for the detailed urban transport plannings in Jakarta before and after the Bus Rapid Transit (BRT) system was commissioned. Each of these studies on gender-heterogeneous commuting time suggest the importance of examining the heterogeneous impact of public transportation on employment by gender, rather than simply an overall effect. More closely related studies have documented the correlations between the access to transportation and labor market outcomes such as income or employment in developing countries (Hyodo et al. 2005; Goel and Tiwari 2016; Glick 1999). These studies use cross-sectional data, so we decide to further extend this line of research by utilizing panel data. A similar line of research using panel data from Lima, Peru on BRT and light rail system is summarized in a working paper by Martínez et al. (2018). But the most relevant analysis, which is ongoing, can be found in the field-experiments being conducted in Lahore, Pakistan for assessing the impact of providing women-only-wagons (a safety measure) to feed into a BRT system on female employment (Majid et al. 2018).[2]

In this paper, we analyze the effects of the Delhi Metro, one of the largest mass rapid transit systems in the current world that has been developed since the early 2000s, on the work participation of women and men, to provide quantitative evidence on whether a high quality urban public transportation system contributes to an improvement in female economic participation. We focus on the Delhi Metro for three reasons. Firstly, Delhi is

2. Majid et al. (2018) reports the effects of BRT on congestion, and a progress of the RCT based impact analysis of safe commuting for female is available in the J-PAL's website: https://www.povertyactionlab.org/evaluation/impact-public-transport-labor-market-outcomes-pakistan

one of the cities in the world fighting against severe concerns for female safety in public spaces and transportations (Jogori and UN Women 2011; Safetipin 2016). According to a Thomson Reuters Foundation Annual Poll in 2017, “New Delhi, the world’s second most populous city with an estimated 26.5 million people, was ranked as the worst megacity for sexual violence and harassment of women alongside Brazil’s Sao Paulo."[3] Also, an UN Women supported survey in Delhi shows that 95 percent of women and girls feel unsafe in public spaces in their 2013 report. Even after the introduction of the Delhi Metro, the situation is still severe but it was even worse before. Recent studies reveal that safety matters to females that have choices in their lives. For example, Borker (2017) finds that safety of school-commuting route has a direct impact on the university choice among the female students in the city of Delhi. In her study, she finds that the willingness to pay for women for a school-commuting route that is one standard deviation safer is an additional 18,800 rupees (290 USD) per year, relative to men, which is an amount equal to double the average annual college tuition. Secondly, India faces challenges over female economic participation and empowerment. Female non-agricultural labor participation has been historically stagnant in South Asia, and there has even been a declining trend in India at the national level (Klasen and Pieters 2015; Andres et al. 2017). For the city of Delhi, while the labor participation of women has not declined, its growth has been stagnating compared to that of men. Lastly, Delhi Metro is one of the best cases to analyze the impact of high quality urban transport infrastructure in developing countries, given its reputations for high service standards. This reputation is not only for its stability and convenience, but also for the safety and comfortable travel of its female passengers. Based on the interviews with users, the introduction of Delhi Metro is shown to have drastically changed transportation choice for women, due to the high standard of safety in the Metro

3. https://poll2017.trust.org/

system (Takaki and Hayashi 2012; Onishi 2017). Motivated by these factors, the existence of the female mobility issue, concerns for female labor supply, and a suitable treatment, we hypothesize the introduction of a safe mode of public transportation in Delhi would have had a non-negligible effect on the supply of female labor (the commuting-safety hypothesis), along with other factors, such as residential relocation, compositional change in labour demand and/or family-level joint labor supply decisions. In this study, we try to quantify the gender-heterogeneous effects of the Delhi Metro system on work-participation rates as the first step in our analysis, solely due to the data limitation.

While our aim has a great policy relevance, it is a difficult research question to obtain a rigorous quantitative answer on because of severe data limitations. First, the standard identification concerns from the non-random location of physical infrastructure are inevitably applicable. This fundamental identification challenge cannot be resolved even if there will be more detailed data available except when there is a suitable natural experiment. Moreover, other impeding facts, like the lack of appropriate individual-level data that covers the period before and after the commission of the Metro as well as the fact that a long time has past since the initial commission of the Metro in 2002, keep us away from making a rigorous causal arguments in an ideal empirical setting.

Our strategy is therefore to use the best-available data and carefully argue its empirical limitations. More specifically, we use the Primary Census Abstract (PCA) which provides various tabulations from the Population Census data for finely disaggregated geographical areas within the National Capital Territory (NCT) of Delhi. We construct a panel of PCA zones for three consecutive census years, 1991, 2001, and 2011. As the measure of intervention, we calculate an accessibility from each PCA zone to the nearest metro station, using maps of PCA zones and the alignment of the Delhi Metro. With the calculated treatment variable, proximity to the Delhi metro, we conduct a difference-in-differences (DID) analysis, controlling for location fixed effect (time-invariant unobserved heterogeneity), to assess whether the proximity to metro stations contributes to the area's growth in female and male participation in non-agricultural economic activities. Since we construct these panel data at the level of the PCA zone-level geographical unit for three rounds (1991, 2001, and 2011) with two pre-treatment periods, we can examine the parallel trend hypothesis which is the prerequisite for DID, by including the "lead" term in the estimation equation.

We find that the effect of the proximity to the Delhi Metro on female work participation rate is positive, and that the same does not seem to hold for men (rather the opposite). This is suggestive evidence that there could be a gender-heterogeneous impact from the Delhi Metro system on the decision of economic participation. In other words, women might respond more positively than men to the proximity to the Delhi Metro stations in deciding whether or not to work.

To understand these empirical findings, we develop a spatial model of urban transportation and commuting. We explicitly model the commuting choice of female and male urban residents who face different commuting costs (fees and travel time plus safety-related welfare cost). We study the model's comparative statics to see how a hypothetical Metro project would affect female and male work participation rates across different zones in a city. We find that if the Metro reduces female commuting costs more than men's, female WPR increases in zones closer to the Metro despite male WPR exhibiting a more ambiguous (or opposite) relationship. This theoretical example shows consistent patterns with our empirical results.

Our empirical findings have a limitation, however, in that the rigorous causal identification of the impact or investigation of a mechanism is affected by the nature and extent of the available data. For example, the gender wage gap or gender-heterogenous comparative advantage in specific skills may result in higher demand for female workers rather than male workers near the metro. However, we do not have gender-specific wage data or skill-level employment information by gender at such a fine geographical unit, so these hypotheses are currently unable to be separated from the commuting-safety hypothesis. Nevertheless, our study is one of the first attempts to quantitatively measure the gendered implication of a large scale urban public transport development in the context of megacities in developing countries.

The rest of the paper is organized as follows. In Section 2, we briefly go over the background of the Delhi Metro project. Section 3 describes the data and Section 4 discusses empirical specifications. Section 5 reports the results. In Section 6, we develop a spatial urban model that shows that the commuting-safety hypothesis has an equilibrium that is consistent with our empirical findings. Section 7 discusses the limitation of our method and potential directions for future research.

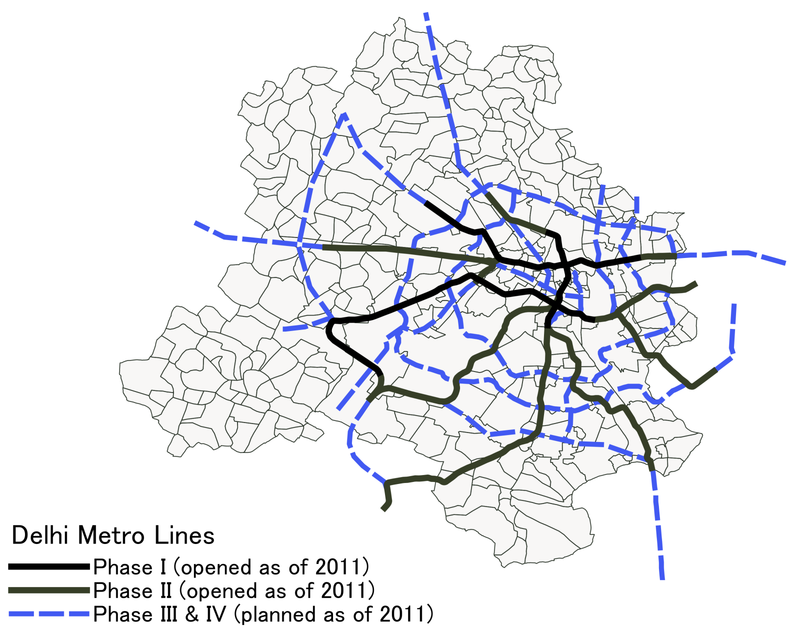
## 2. Background of Delhi Metro

As the country's third urban mass rapid transit system (MRT) and the first of its kind in the capital city, the Delhi metro project has been developed over the past seventeen years. The first phase of Delhi Metro project consisted of the 58 stations and lines covering 65km and commissioned during 2002-2006. Following the Phase I, Phase II built 85 stations and lines covering 125km, which were commissioned during 2008-2011. As of the end of 2011, Phase III and Phase IV were in the planning stage. The geographical alignments of the Delhi Metro lines in the different phases are shown in Figure 1.

The novelty of the Delhi Metro project is the fact that it focused on the safety and inclusiveness from its planning stage. Adaptation of women-only car, barrier-free design, rubbish control for keeping train clean, and security check at the entry have contributed towards providing safe public mass urban transportation for the citizens of Delhi. Overall, the Delhi Metro has gained a reputation for high standard of facility and operation that ensures safety and comfort for female passengers (Takaki and Hayashi 2012; Onishi 2017).

Prior to the introduction of the Delhi Metro, safety concerns in the public transportation system had been severe for women in Delhi (Jogori and UN Women 2011; Safetipin 2016). While affordable and reliable urban transportation plays a vital role in engaging in either income-generating activities and schooling in optimal locations, or other activities such as household chores, family visits, or leisure, it is not difficult to hypothesize that the limitation of safe modes of transportation was taxing for women in trying to get access to these social and economic opportunities. Given such a context, the introduction of a relatively safe public transportation system has had a potential impact to drastically change women's behavior in Delhi.

**Figure 1:** Delhi Metro Alignment




*Source:* The authors construct this map based on GIS maps procured from the following GIS map vendors. Base-map with zone boundaries: Zenrin Co., Ltd. Metro Alignment: Compare Infobase Ltd.

## 3. Data

We use the Primary Census Abstract (PCA) of India's Population Census in 1991, 2001, and 2011, published by the Office of the Registrar General and Census Commissioner, Ministry of Home Affairs. The PCA provides aggregates of population census enumeration at the level of a small local administrative unit and/or a ward of constituency. In 1991, entire area of the NCT of Delhi was divided into towns, villages, and charges.[4] Since the geographical boundaries of administrative units change overtime, we interpolate the data of 2001 and 2011 based on area size so that the boundary is consistent with that of 1991. We carry out spatial interpolation as follows. To simplify the explanation, we consider the case of two period, period 0 and period 1. Suppose there are a total of $J_0$ zones in the period 0, indexed as $j_0 = 1, ..., J_0$. In period 1, suppose there are a total of $K_1$ zones indexed as $k_1 = 1, ..., K_1$. The boundaries of zones are not in general consistent between the two periods, which means that a zone in period 0 intersects with multiple zones in period 1. Consider a particular zone $j_0$ of the period 0 which intersects with multiple period 1 zones. Let $S^1_{j_0}$ denote the set of these period 1 zones intersecting with $j_0$. For each of these period 1 zones $k_1 \in S^1_{j_0}$, the area can be divided into a part intersecting with $j_0$, denoted as $a^{j_0}_{k_1}$, and the remaining part, $a^{-j_0}_{k_1}$ which does not intersect with $j_0$. Our spatial interpolation calculates the period 1 value of zone $j_0$ statistics by taking a weighted average of the statistics of the intersecting period 1 zones in $S^1_{j_0}$. More specifically, the interpolated value of variable $x$ for zone $j_0$ in period 1 is given by:

$$\tilde{x}^1_{j_0} = \sum_{k_1 \in S^1_{j_0}} \frac{a^{j_0}_{k_1}}{a^{j_0}_{k_1} + a^{-j_0}_{k_1}} x_{k_1} \tag{1}$$

4. Charge is an electorate unit which disintegrated the central part of Delhi in the 1991 Census. In the 1991 Census, disaggregated data for the MCD (Municipal Corporation of Delhi), consisting the central part of the NCT of Delhi, were reported at the level of Charge.

This interpolation only applies to the variables in levels, such as population and the number of workers. For the variables in rates, we calculate them using the interpolated level variables. For example, a rate variable $r$ which is defined as the ratio of two level variables $x$ and $y$, or $r = \frac{y}{x}$, we obtain the period 1 interpolated value by $\tilde{r}^1_{j_0} = \frac{\tilde{y}^1_{j_0}}{\tilde{x}^1_{j_0}}$. To check the robustness of the key results of this interpolation, we add the analyses using only those zones with consistent boundaries over time in Section C of the appendix.

To represent the economic participation of each gender group from the available statistics, we calculate "(non-agricultural) work participation rate" ("WPR" hereafter). The work participation rate is measured by the ratio of the number of "main workers" (works more than 6 months per year) in "other sectors"(other than cultivators, agricultural labourers, or household industry workers)[5] divided by the adult population[6], for each gender. This indicator is different from the labor force participation rate (LFPR). While the denominator of LFPR is usually the working-age population above the age of 15, the denominator of WPR is the (imputed) adult population. Moreover, the numerator is also different because the definition of being a labor force includes those who are employed and unemployed, while that of work participation rate does not include those who are seeking for a job. These definitional differences make WPR either smaller or larger than LFPR, which is an empirical question because the difference in the denominators depends on how all-ages population minus two times the 0-6 population differs from the population over 15. In fact, the urban areas' LFPR during these periods has increased from 14.7 percent to 15.5

5. The "Other Sector": All workers, i.e., those who have been engaged in some economic activity during the last one year, but are not cultivators or agricultural labourers or in the Household Industry, are 'Other Workers(OW)'. The type of workers that come under this category of 'OW' include all government servants, municipal employees, teachers, factory workers, plantation workers, those engaged in trade, commerce, business, transport banking, mining, construction, political or social work, priests, entertainment artists, etc. In effect, all those workers other than cultivators or agricultural labourers or household industry workers, are "Other Workers".

6. Since the adult population is not given in PCA, we impute it by "total population - 2 x (population of 0 to 6 ages)", base on the population pyramid of India.

percent, while Delhi's WPR increased from 7.06 percent to 7.91 percent. Though the levels are different due to the definitional difference discussed above, the trend is consistent across the two measures.

Our treatment variable is the proximity of a zone (a town, a village, or a charge based on the 1991 administrative boundaries) to its nearest Metro Phase I and II stations. To represent the proximity to Metro stations, we measure the average distance using the coordinates of boundaries of towns and villages, as well as the alignment of the Metro stations. The average distance measure is constructed as follows: (i) A large number of equally spaced points (about 0.5 million) are generated and plotted over the entire area of Delhi; (ii) From each point, the nearest Metro station is searched and the distance from the point to the nearest Metro station is calculated. For a point $k$ located within the boundary of zone $i$, this distance is denoted as $d_{k(i)}$; and (iii) The average distance to the nearest Metro station(s) of the zone $i$, $D_i$, is then calculated as

$$D_i = \frac{\sum_{k(i)} d_{k(i)}}{N_i} \tag{2}$$

where, $N_i$ is the number of points in zone $i$. $D_i$ is smaller (i.e. the treatment intensity is larger) if $i$ is closely located to Metro stations opened during 2002-2011, after the commission of the Phase I and II Metro network. Average distance measures to the the Metro Phase III and IV (only under the planning phase in 2011) are also calculated in the same manner to better define the comparison group that is more likely to share the similar unobserved characteristics regardless of the assigned treatment. In addition, we use the total population, the number of children, the number of households, the number of literal residents, and the number of residents scheduled caste (each by gender) from the PCA tables as control variables in the main analysis (the analysis without these controls are available in the robustness check).[7]

The descriptive statistics is shown in the Table 1. The upper table summarises our time-invariant variables, and the lower one is for time-variant variables. Our time-invariant variables are the distances to the Phase I and II metro stations that had commissioned by 2011, and those of the Phase III and IV which had not yet opened. On average, distance to the nearest Phase I or Phase II metro station is 5.2 km. Since the location of the planned metro stations, those of Phase III and IV, are more stretched out to the suburbs, the average distance to the Metro station is shorter 3.3km.

The time-variant variables are the outcome variables and control variables used in the estimation. Female WPR has been substantially lower than that of males throughout two decades since 1991. However, their average WPR has increased from 5.3 percent in 1991 to 7.9 percent in 2011, while men's WPR has grown from 40.4% to 45.3% during the same period. In contrast to the national level decline in labor force participation of women, the small increase of their WPR in Delhi might be partially due to the contribution of the Delhi Metro.

Figure 2 depicts kernel density estimates for the distribution of female and male WPRs for years 1991, 2001, and 2011. First, we can observe that the WPR distributions are distinctly different between the two gender groups in each year. That of females are clustered at lower rate of WPR with smaller variance, in contrast to that of males. Secondly, there is a subtle, but universal shifts of female WPR distribution towards the right. This suggests that the rate was improving almost everywhere in the distribution for women. Male WPR initially had a flat distribution in 1991 while it had evolved into a single peaked one in 2001. We do not observe a distinct shift in the distribution from 2001 to

7. Under the constitution of India, scheduled caste is defined as follows. `http://socialjustice.nic.in/writereaddata/UploadFile/Compendium-2016.pdf`. India's Census follows this definition. See `http://censusindia.gov.in/Census_And_You/scheduled_castes_and_sceduled_tribes.aspx`.

2011.[8]

Figure 3 shows the spatial distribution of WPR of females and males for the two census years, 2001 and 2011. The dark-red zones are places with the highest WPR and the dark-blue zones are with the lowest WPR. The top two panels, 3a and 3b show women's WPR. The bottom two, 3c and 3d are those for men.

**Table 1:** Summary Statistics

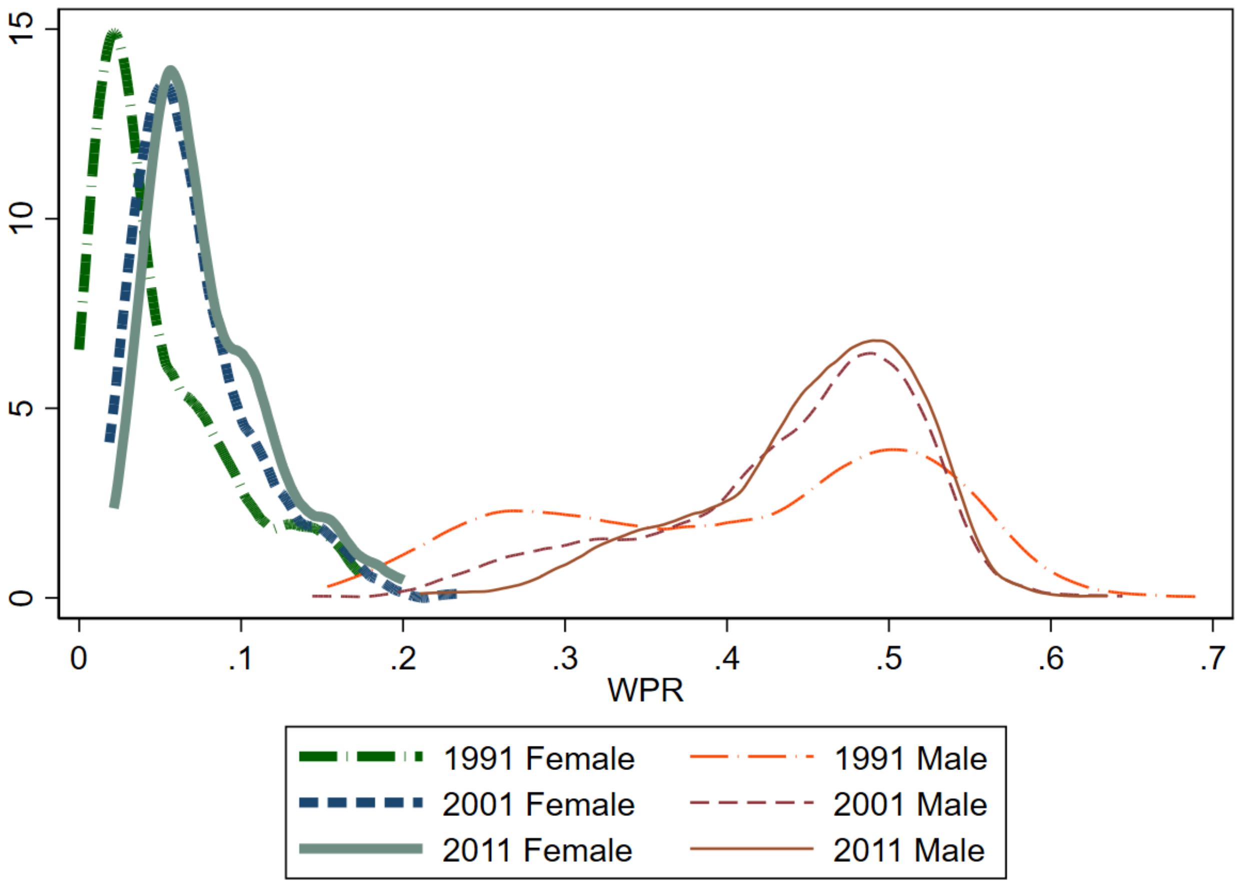
| Time-invariant variables | (1) N | (2) mean | (3) sd |
|---|---|---|---|
| Dist. to Phae I or II Metro St. (km) | 342 | 5.239 | 4.763 |
| Dist. to Phae III or IV Metro St. (km, used for sub-sample selection) | 342 | 3.274 | 3.145 |

| | 1991 | | | 2001 | | | 2011 | | |
|---|---|---|---|---|---|---|---|---|---|
| Time-variant variables | (1) N | (2) mean | (3) sd | (4) N | (5) mean | (6) sd | (7) N | (8) mean | (9) sd |
| Female WPR | 332 | 0.0531 | 0.0599 | 342 | 0.0706 | 0.0369 | 342 | 0.0791 | 0.0371 |
| Male WPR | 332 | 0.404 | 0.131 | 342 | 0.439 | 0.0812 | 342 | 0.453 | 0.0674 |
| Female to male WPR ratio | 332 | 0.118 | 0.0993 | 342 | 0.161 | 0.0864 | 342 | 0.171 | 0.0644 |
| Household Size | 332 | 5.562 | 0.982 | 342 | 5.283 | 0.478 | 342 | 5.038 | 0.396 |
| Children Share | 332 | 0.184 | 0.0369 | 342 | 0.150 | 0.0235 | 342 | 0.124 | 0.0171 |
| Female to male literacy ratio | 332 | 0.698 | 0.163 | 342 | 0.817 | 0.0679 | 342 | 0.865 | 0.0485 |
| Female to male SC ratio | 327 | 1.007 | 0.155 | 342 | 1.042 | 0.0574 | 342 | 1.027 | 0.0362 |

*Note:* The upper table summarizes the time-invariant variables. The lower one is for the time-variant variables.

8. The T-test comparing the means of WPR across different years show that the mean WPR has significantly increased over time for both genders. We also conduct the Kormogolov-Smirnov test to statistically assess whether the WPR distribution changes across years. The results indicate that women's WPR has different distribution across three census years, while the men's distribution between 2001 and 2011 are not statistically distinguishable. The results are reported in Table A.1 in the Appendix.

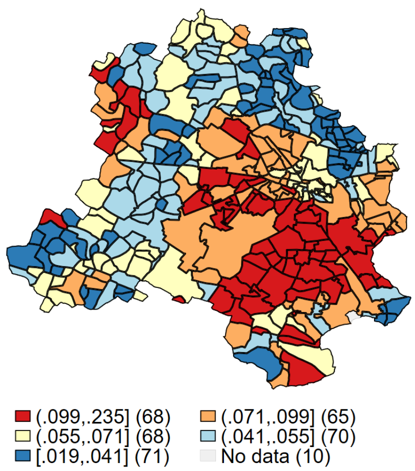
**Figure 2:** Kernel Distribution of Female and Male WPR, 2001 and 2011



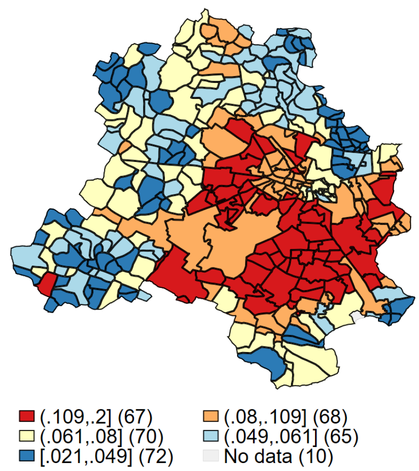

*Source:* Authors

**Figure 3:** Spatial Distribution of WPR for females and males, in 2001 and 2011

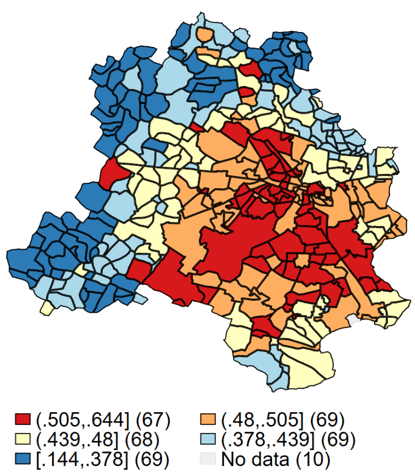
(a) 2001 Female WPR



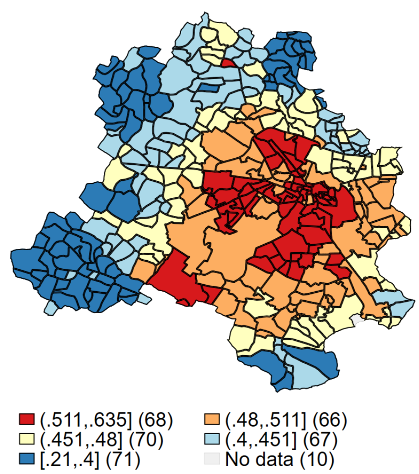

(b) 2011 Female WPR




(c) 2001 Male WPR




(d) 2011 Male WPR

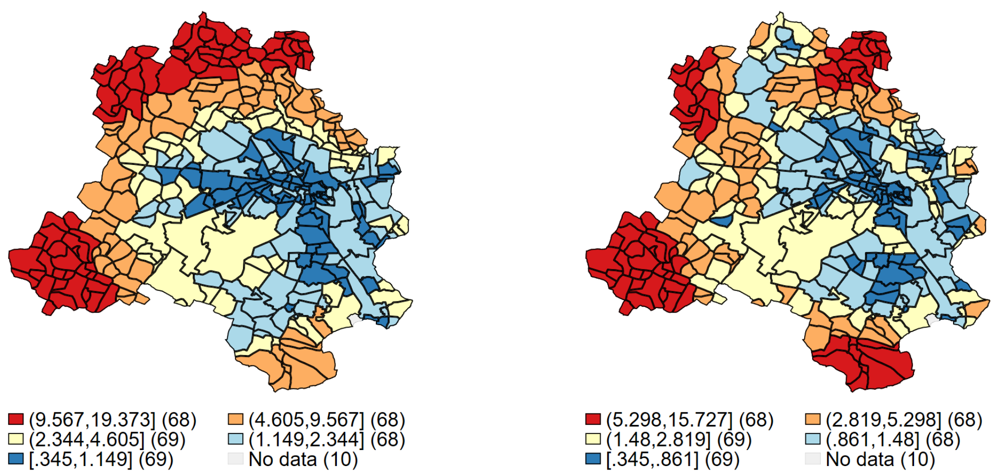



*Source:* Authors

**Figure 4:** Distance to Commissioned and Planned Metro Stations

(a) Distance to Commissioned (PH I & II) Metro Stations

(b) Distance to Planned (PH III & IV) Metro Stations




*Source:* Authors

# 4. Empirical Strategy

As described in Section 3, our data is neither experimental nor quasi-experimental. The unit of observation is aggregated at the level of zones (town or ward), which divide the NCT (National Capital Territory) of Delhi into 342 geographical units. Using a panel data of zones in Delhi for 1991, 2001 and 2011, we employ the difference-in-difference (DID) method with two pre-treatment periods (1991, 2001). In estimation, we check for parallel pre-trend by exploiting these two pre-treatment periods. Specifically, we estimate the following equation:

$$Y_{it} = \beta_0 + \beta_{-1}(D_i \times Pre_t) + \beta(D_i \times Post_t) + \delta X_{it} + \theta_t + \alpha_i + \epsilon_{it}, \tag{3}$$

Where, $Y_{it}$ is the outcome variable of zone $i$ at year $t$; $D_i$ is the treatment variable, the log of average distance to nearest Phase I or II metro station, $Pre_t$ is a time dummy taking 1 when $t = 1991$ and 0 otherwise. $Post_t$ is another time dummy taking 1 if $t = 2011$ and 0 otherwise. Coefficient $\beta$ is our central interest. This is the post-treatment effect of the distance to the metro station on the outcome. $\beta_{-1}$ captures the correlation between a zone's distance to the metro station and the outcome variables in 1991. If $\beta_{-1}$ is insignificant, we will not reject the hypothesis that pre-trend is not associated with the distance to a metro station, suggesting that the parallel pre-trend hypothesis holds. The same strategy has been used in studies such as Autor (2003) and Kearney and Levine (2015). $X_{it}$ is a vector including other time-variant location specific characteristics such as average household size, share of children (under 6 years old) in the population, female literacy rate relative to male, and ratio of share of scheduled caste between female and male.[9] The first two variables are introduced to control for the variations in the presence of dependents in household (i.e. elderly and children) which are not directly measured in the PCA. The latter two control for the variation in the gender inequality.[10] Lastly, $\theta_t$ is year fixed effect, $\alpha_i$ is a zone-level fixed effect, and $\epsilon_{it}$ is the error term.

The goal of this paper is to empirically assess how the Delhi Metro differently affects the economic participation of women and men. More specifically, we investigate whether the zones closer to the Delhi Metro station have observed more increase in work participation of the residents than those in zones further away from metro stations, separately for women and men. In the empirical analyses below, we focus on four measures of work participation, female WPR, male WPR, a ratio of a zone's female WPR to male WPR ($= WPR(women)/WPR(men)$), or "WPR ratio" in short; and WPR for total residents (sum of females and males).

For the treatment variable, $D_j$, we define the (log) distance to the nearest Phase I or II Delhi Metro station. The reason for this choice of continuous treatment variable is twofold. Firstly, we would like to avoid a discretionary construction of a treatment variable, which is unavoidable when using discrete variables (i.e., we do not know from which kilometer it is "proximate" to the metro). Secondly, it is rather easier to interpret the results.

We estimate this equation 3 with a standard fixed effects estimator. The coefficient $\beta$ will capture the treatment effect, and the sign and the magnitude of this coefficient is our central concern. $\beta_{-1}$ is the coefficient on the "lead" term. We expect that $\beta_{-1}$ is insignificant under the common trend assumption. Please note that the insignificance of

9. For clarity, variables are given by; average household size $= \frac{\text{Population}}{\text{Number of Household}}$ ; share of children (under 6 years old) in the population $= \frac{\text{Number of Children (under 6)}}{\text{Population}}$; female literacy rate relative to male $= \frac{\text{female literacy rate}}{\text{male literacy rate}}$; and ratio of share of scheduled caste between female and male $= \frac{\text{share of scheduled caste in female population}}{\text{share of scheduled caste in male population}}$

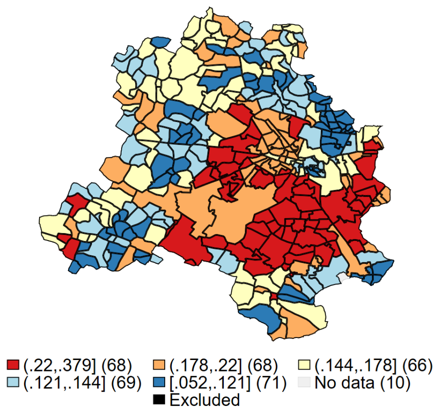
10. In the separate regression, we check that these variables do not seem to be the consequences of the treatment $D_i \times Post_t$, allowing us to included them as controls in the equation.

$\beta_{-1}$, is only suggestive evidence that the two sets of zones (in this case near and far from the new metro stations) would have evolved similarly in the absence of the intervention. It is not decisive as to whether there is unobserved heterogeneity across regions affecting the change in outcome variables or not. In fact, recent studies such as Kahn-Lang and Lang (2019) as well as Jaeger, Joyce, and Kaestner (2018) note that the parallel pre-trends do not necessarily imply parallel trends.

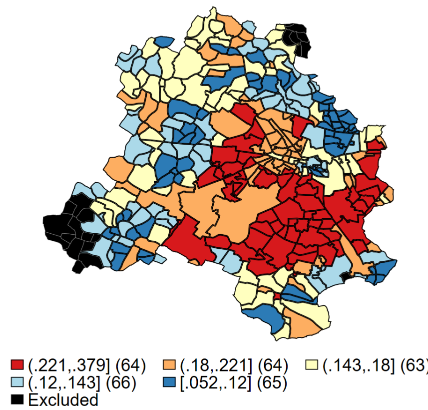
We conducted the estimation across various sub-samples to see how the results are sensitive to the selection of the comparison group. We compare five sub-sample defined as follows; (1) All the zones in Delhi (Figure 5a); (2) includes only the zones within 10km reach from the nearest commissioned (Phase I or II) station or the nearest planned (Phase III or IV) Metro station (Figure 5b); (3) includes only the zones within 5km reach from the nearest commissioned (Phase I or II) station or the nearest planned (Phase III or IV) Metro station (Figure 5c); (4) trims the zones in the subset (2) so that it include only zones at least 10km further from the CBD of Delhi, Connaught Place (Figure 5d); and (5) trims the zones in the subset (3) so that it include only zones at least 10km further from the CBD of Delhi, Connaught Place (Figure 5e).

**Figure 5:** Subsample Definition and "WPR ratio" in 2011

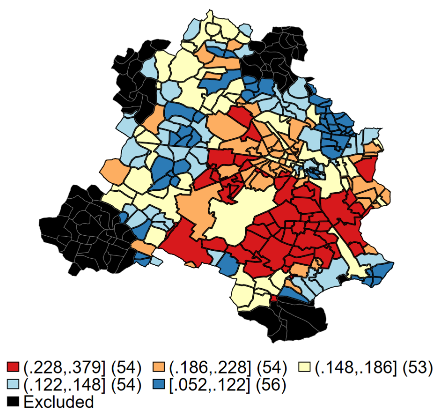
(a) All zones in Delhi (1)



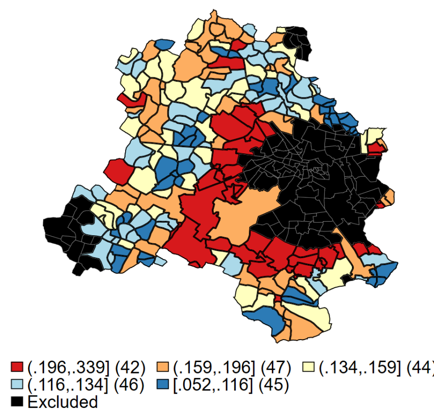

(b) Within 10km reach from commissioned and planned Metro Stations (2)



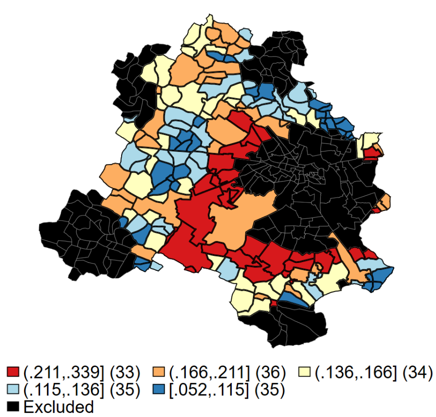

(c) Within 5km reach from commissioned and planned Metro Stations (3)




(d) Within 10km reach from commissioned and planned Metro Stations and at least 10km further from the CBD (4)




(e) Within 5km reach from commissioned and planned Metro Stations and at least 10km further from the CBD (5)




*Source:* Authors

## 5. Results

Tables 2-5, report the results of the estimations across different specifications. Table 2 reports the estimation results of equation (3) taking the female WPR as the outcome. For all the five subset analyses, our treatment variable, $D_{it}$, is significant at the 1 percent significance level with negative sings, except for column (5) where significance is at the 5 percent significance level. Negative coefficient indicates that being close to the commissioned Metro station makes female work participation rate higher. For example, for the full sample case, shown in column (1) of the Table 2, if the distance to the nearest Phase I or II station becomes doubles, female WPR decreases by 0.558 percentage points. Given that the mean of female WPR in 2011 was 7.91 percent, this implies that doubling the distance around the mean distance of 5.239km will reduce WPR of females to 7.35 percent.

Columns (2) and column (3) of Table 2 limit the sample zones to within 10km and 5km access to any Metro station regardless of whether they had already been commissioned as of 2011 (i.e., "control group" is restricted to the areas near Phase III or IV). We regard that the zones closer to the planned network are “selected” for Delhi Metro intervention, but the metro service is not yet available at that point in time, so they may share the similar unobserved heterogeneity with zones close to the commissioned stations, which affect the change in outcomes.[11] By estimating the model of columns (2) and (3), we compare the outcomes in zones for those who got access to metro stations earlier with those who would get it later. Therefore, by estimating the model only in those areas close to either the commissioned or planned metro stations, we compare outcomes in zones between the areas gaining access to the metro stations earlier (before 2011) and those would gain access later

11. To identify the causal impact of transport system, it is common to use planned but never developed routes as control group; however, there is no such locations in Delhi Metro’s case. Instead, we adopt an idea close to the phase-in approach for improving our identification. The phase-in approach in our context (or in transportation infrastructure projects in general) still suffers from the remaining endogeneity bias due to the non-random construction timing/order of projects.

(post 2011).

We also note that the effect seems to be stronger outside the central area. The magnitude of the coefficient is greater for column (4), the outer area subsample, than that of column (2) (the cut-off at 10 km). The same argument applies to the comparison between columns (3) and (5), where the cut-off is 5km.

The results shown in the Table 2 suggest that a positive effect of the accessibility to the Delhi Metro for females exists throughout all the specifications. Furthermore, for all columns, the coefficients on the "lead" term are insignificant, which means that the common pre-trend assumption holds for these sets of analyses.

Table 3 shows the results for the effects on male WPR. Contrary to the case for females, all the coefficients on the distance to a commissioned Metro station are positive and significant at the 1 or 5 percent significance level. The parallel pre-treatment trend assumption is overall satisfied except for column (3) whose coefficient $\beta_{-1}$ term is negative and statistically significant at the 10 percent level. Furthermore, the magnitude of the effect does not vary across subsamples, ranging from 0.00801 to 0.00975, compared to the case for females as shown in Table 2. From the results in Table 2 and Table 3, it turns out that the proximity to the Delhi Metro station affects positively for female WPR while its effects is negative for that of male. Given that the mean of WPR for males in 2011 is 45.3%, this implies that doubling the distance around the mean distance of 5.239km will increase the WPR of males to 46.2%.

Table 4 reports the results when the outcome variable is the WPR ratio between female and male. Consistent with the results in Table 2 and Table 3, the coefficients on the distance to commissioned station are negative and significant at the 1 percent level. The results implies that the gap of WPR between females and males becomes slightly smaller (i.e. WPR ratio increases) in zones closer to commissioned Metro station. The key identifying assumption is again the common trend, and it seems to be satisfied for the trend between 1991 and 2001 as the coefficient $\beta_{-1}$ is insignificant.

Finally, Table 5 reports the results when the total WPR is used as the outcome. Total WPR is the sum of female and male main workers in the non-agricultural sector divided by total adult population. For the first three columns show significantly positive coefficients on the distance to the nearest commissioned Metro station, meaning that proximity to a Metro station has a negative effect on total work participation. However, as shown in columns (4) and (5), the effect becomes no longer significant in suburban subsamples. This is mainly due to the imprecision in smaller sample sizes as the point estimates do not change in magnitude comparing to columns (1)-(3), whiles the standard errors are larger. In the area outside of the CBD premises, proximity to the Metro does not change the overall work participation.

We add several robustness checks to address a series of technical concerns. The first relates to the control variables we included in the estimation equation. If the control variables are endogenous to the treatment variable, the inclusion of the controls in the equation is problematic (Angrist and Pischke 2009). Therefore, we conduct the same estimations without the control variables, as reported in Tables B.2-B.5 of Section B of the Appendix. For women, the results are qualitatively the same as our main estimation. For men, we have qualitatively similar results for our variable of interest ("Dist. to Metro (2011)") without control variables compared with the case of our main model reported in Table 3. However, the $\beta_{-1}$ term becomes highly significant in this case. This implies that the control variables we include in the main analyses capture the pre-trend heterogeneity for the case of male WPR well.

We also check the sensitivity of the results against our method of interpolating the data so that the boundary definition of the "zones" would be consistent with that of 1991. In

Appendix C, we introduce an illustrative explanation of our interpolation method and its potential effects on the statistical outcomes. We also report the estimation results only using the data of zones with consistent boundaries. Again, the female results are stable, while the male ones are sensitive to the choice of sample zones.

From the empirical findings above, we can summarize the effects of the Delhi Metro on the work participation as follows. Firstly, female WPR in 2011 is higher in zones close to the Delhi Metro station, while it is lower in the more distant zones whereas male WPR is instead higher in these zones. From the results of robustness checks, estimating the equation without controls and using only the 1991-boundary consistent subsample, we find that the results for male WPR are sensitive to the settings while female ones are stable overall. Therefore, it is safe to conclude that in the areas closer to Metro, the economic participation of women expanded more intensively than that of men. Secondly, for women, the magnitude of the effects is larger in the suburban subsamples. This means that the heterogeneous responses by gender caused by the access to the Delhi Metro might be more pronounced in the suburban area than in the CBD. Thirdly, partially reflecting the fact that women are positively affected by proximity and men are negatively affected, the total WPR is negatively affected, because the effects on men surpass those on women. The last point suggests that it is important to separately analyze the effects of transportation between females and males without averaging out the overall effects.

What is the mechanism that delivers this gender differentiated outcomes? One potential story could be an additional mobility benefit that the Delhi Metro may provide for women. The Delhi Metro is a mass rapid transit system which did not exist before in that city, where road vehicles such as buses, three-wheelers, and rickshaws, are the major mode of the transportation. The Delhi Metro has given citizens a faster and more reliable (predictable) method for travel in the city. The time-saving effect of the Metro system contributes to a reduction of the travel cost of both males and females who use the system. Before the introduction of the Metro, it is a plausible conjecture that female mobility was substantially more constrained than men, considering the safety issues including sexual violence on public transport. If the Delhi Metro secures a mode that allows females to travel more safely than on other traditional modes, the effective travel cost for might be reduced by more than just the time-saving effect. If this additional benefit is large, then whether to live in the neighbourhood of the Metro station should matters more to the mobility of females than to that of males. To argue this implication more formally, Section 6 introduces a theoretical model that explains this potential mechanism.

Other than that, there are a couple of other mechanisms that could generate gender-heterogeneous effects. Firstly, labor demand might change by the introduction of the Metro and that could be gender-heterogenous. For example, the manufacturing sector might relocate its factories and offices outside of the downtown core, while service sector jobs might flourish downtown. Each sector might attract workers of a different gender. One possibility is that males are mainly in manufacturing, and this causes their residences to move further away from the metro, which is rather concentrated near the downtown core, as the factories relocate to the suburbs. An alternative possibility is that women tend to take job openings in service sector, and residents near the metro start working as it stimulates demand for service sector jobs. Secondly, a reduction of congestion and travel time that can plausibly benefit both females and males but in different magnitudes, encourages the residents to commute further. Nevertheless, it also induces in-migration of workers into areas near the Metro stations, and this results in higher work participation rates and housing prices in those areas.[12] The resulting residential relocation itself is hard

12. How WPR and housing price react also depends on the elasticity of housing supply, the spatial allocation of industries within cities, and wage and many other things, making the actual signs and magnitude of the impact ambiguous.

to analyze due to the data limitations, but the imapct through this channel could be gender-heterogenous as well. Thirdly, it is also important to note that the family level decision process can complicate male and female labor supply decisions. For example, if a family (couple) faces a reduction of travel time by the Metro and a high paid job becomes available to the husband, one of the possible responses is the wife's withdrawal from labor market activity (increase home production), substituting for this an increase in the male labor supply (i.e., intensification of division of labor). When all of these effects are combined, it is possible that the family with a male-bread winner moved away from the metro into more reasonably priced residential areas and his wife resigns job or does not seek employment. This story, however, cannot fully explain a subtle but statistically significant gain in female WPR closer to the metro stations, so we still think our safe-commuting hypothesis will survive the further tests in future research. The remaining challenge would be to separately identify the safe-commuting hypothesis versus gender-heterogeneous shifts in labor demand (with crowd-out and/or location segregation by industry-gender combination) because both stories can explain the current empirical findings on females.

**Table 2:** Effects of Proximity to the Delhi Metro on the Female Work Participation Rate (Difference-in-Differences)

| VARIABLES | (1)<br>ALL | (2)<br>d < 10km | (3)<br>d < 5km | (4)<br>d < 10km<br>& CBD > 10km | (5)<br>d < 5km<br>& CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | -0.00558*** | -0.00688*** | -0.00418*** | -0.00906*** | -0.00453** |
| | (0.00146) | (0.00166) | (0.00152) | (0.00230) | (0.00207) |
| Dist. to Metro(1991, $\beta_{-1}$) | 0.00186 | 0.00173 | 0.00505 | 0.00152 | 0.00479 |
| | (0.00316) | (0.00355) | (0.00376) | (0.00362) | (0.00430) |
| Household Size | -0.0855*** | -0.0852** | -0.0810** | -0.0898* | -0.0797* |
| | (0.0317) | (0.0344) | (0.0317) | (0.0463) | (0.0439) |
| Children Share | -0.124*** | -0.131*** | -0.136*** | -0.113*** | -0.129*** |
| | (0.0215) | (0.0226) | (0.0256) | (0.0222) | (0.0263) |
| Female to male literacy ratio | -0.0923** | -0.0938** | -0.0589 | -0.123*** | -0.0911** |
| | (0.0396) | (0.0397) | (0.0419) | (0.0330) | (0.0376) |
| Female to male SC ratio | -0.0629*** | -0.0727*** | -0.0454* | -0.0831*** | -0.0553* |
| | (0.0214) | (0.0256) | (0.0261) | (0.0295) | (0.0295) |
| Constant | -0.0402 | -0.0519 | -0.0636 | -0.0192 | -0.0593 |
| | (0.0650) | (0.0690) | (0.0758) | (0.0739) | (0.0804) |
| Year Dummy | YES | YES | YES | YES | YES |
| Observations | 1,006 | 948 | 801 | 654 | 507 |
| R-squared | 0.443 | 0.449 | 0.431 | 0.529 | 0.470 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Adj-R | 0.438 | 0.444 | 0.426 | 0.523 | 0.462 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
“d < km” if sample zones with distance to Phase I - IV stations within x km
“CBD < km” if sample zones locate further than x km from the CBD

**Table 3:** Effects of Proximity to the Delhi Metro on the Male Work Participation Rate (Difference-in-Differences)

| VARIABLES | (1) ALL | (2) d < 10km | (3) d < 5km | (4) d < 10km & CBD > 10km | (5) d < 5km & CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | 0.00862*** | 0.00993*** | 0.00975*** | 0.00801** | 0.00829** |
| | (0.00217) | (0.00225) | (0.00237) | (0.00327) | (0.00329) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.00638 | -0.00829 | -0.0102* | 0.00160 | -0.00281 |
| | (0.00493) | (0.00525) | (0.00609) | (0.00428) | (0.00485) |
| Household Size | -0.347*** | -0.332*** | -0.334*** | -0.360*** | -0.359*** |
| | (0.0334) | (0.0333) | (0.0330) | (0.0404) | (0.0410) |
| Children Share | -0.0362 | -0.0432 | -0.0719* | -0.00120 | -0.0471 |
| | (0.0412) | (0.0427) | (0.0368) | (0.0460) | (0.0308) |
| Female to male literacy ratio | 0.0268 | 0.0191 | 0.0522 | -0.0329 | -0.0140 |
| | (0.0525) | (0.0522) | (0.0649) | (0.0202) | (0.0263) |
| Female to male SC ratio | 0.0353 | 0.0163 | 0.0451 | 0.0510 | 0.0742 |
| | (0.0433) | (0.0450) | (0.0477) | (0.0460) | (0.0506) |
| Constant | 0.951*** | 0.918*** | 0.883*** | 1.011*** | 0.943*** |
| | (0.103) | (0.104) | (0.0986) | (0.106) | (0.0808) |
| Year Dummy | YES | YES | YES | YES | YES |
| Observations | 1,006 | 948 | 801 | 654 | 507 |
| R-squared | 0.462 | 0.456 | 0.455 | 0.609 | 0.600 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Adj-R | 0.457 | 0.451 | 0.449 | 0.604 | 0.594 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

**Table 4:** Effects of Proximity to the Delhi Metro on a ratio of the Work Participation Rate of Females over that of Males (Difference-in-Differences)

| VARIABLES | (1)<br>ALL | (2)<br>d < 10km | (3)<br>d < 5km | (4)<br>d < 10km<br>& CBD > 10km | (5)<br>d < 5km<br>& CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | -0.0166*** | -0.0210*** | -0.0120*** | -0.0278*** | -0.0160*** |
| | (0.00438) | (0.00476) | (0.00395) | (0.00676) | (0.00542) |
| Dist. to Metro(1991, $\beta_{-1}$) | 0.000604 | 0.00132 | 0.0120 | -0.00377 | 0.0101 |
| | (0.00674) | (0.00748) | (0.00760) | (0.00919) | (0.00969) |
| Household Size | -0.138** | -0.145** | -0.131** | -0.164** | -0.134* |
| | (0.0556) | (0.0598) | (0.0546) | (0.0809) | (0.0753) |
| Children Share | -0.265*** | -0.279*** | -0.230*** | -0.280*** | -0.219*** |
| | (0.0412) | (0.0411) | (0.0410) | (0.0532) | (0.0459) |
| Female to male literacy ratio | -0.130** | -0.128** | -0.0700 | -0.172*** | -0.117** |
| | (0.0567) | (0.0569) | (0.0614) | (0.0466) | (0.0541) |
| Female to male SC ratio | -0.153** | -0.131** | -0.0770* | -0.164** | -0.0952* |
| | (0.0614) | (0.0563) | (0.0460) | (0.0730) | (0.0539) |
| Constant | -0.136 | -0.151 | -0.0805 | -0.120 | -0.0575 |
| | (0.117) | (0.120) | (0.120) | (0.149) | (0.135) |
| Year Dummy | YES | YES | YES | YES | YES |
| Observations | 1,006 | 948 | 801 | 654 | 507 |
| R-squared | 0.348 | 0.361 | 0.387 | 0.379 | 0.365 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Adj-R | 0.343 | 0.356 | 0.381 | 0.372 | 0.355 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

**Table 5:** Effects of Proximity to the Delhi Metro on the Work Participation Rate of the Sum of Females and Males (Difference-in-Differences)

| VARIABLES | (1) ALL | (2) d < 10km | (3) d < 5km | (4) d < 10km & CBD > 10km | (5) d < 5km & CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | 0.00326** | 0.00326** | 0.00406** | 0.00131 | 0.00289 |
| | (0.00152) | (0.00163) | (0.00160) | (0.00233) | (0.00218) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.00233 | -0.00328 | -0.00299 | 0.00170 | 0.000285 |
| | (0.00381) | (0.00419) | (0.00475) | (0.00321) | (0.00364) |
| Household Size | -0.254*** | -0.248*** | -0.249*** | -0.266*** | -0.263*** |
| | (0.0264) | (0.0278) | (0.0261) | (0.0359) | (0.0351) |
| Children Share | -0.0787*** | -0.0859*** | -0.103*** | -0.0578* | -0.0900*** |
| | (0.0295) | (0.0309) | (0.0284) | (0.0317) | (0.0232) |
| Female to male literacy ratio | -0.0296 | -0.0339 | -0.000773 | -0.0773*** | -0.0543** |
| | (0.0444) | (0.0443) | (0.0518) | (0.0210) | (0.0229) |
| Female to male SC ratio | 0.0118 | 0.00323 | 0.0314 | 0.0116 | 0.0355 |
| | (0.0295) | (0.0296) | (0.0315) | (0.0281) | (0.0297) |
| Constant | 0.539*** | 0.519*** | 0.499*** | 0.582*** | 0.531*** |
| | (0.0757) | (0.0785) | (0.0778) | (0.0779) | (0.0648) |
| Year Dummy | YES | YES | YES | YES | YES |
| Observations | 1,006 | 948 | 801 | 654 | 507 |
| R-squared | 0.474 | 0.466 | 0.467 | 0.642 | 0.634 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Adj-R | 0.469 | 0.461 | 0.461 | 0.638 | 0.629 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

# 6. Theoretical Explanation with a Spatial Commuting Model

The empirical results indicate that the commission of the Delhi Metro raises the work participation rate of women living nearby the Metro stations, while its effect is ambiguous for men (this is the opposite of women in our main specification, but it is highly sensitive to the inclusion of control variables).

In what follows, we try to argue that the observed results in WPR for women and men can be caused by the heterogeneous reduction of commuting cost by gender, owing to the Metro. Here, the commuting cost includes not only fees and opportunity cost, but also a safety-related welfare cost. Especially, using a simple theoretical model of urban spatial economy, we find that women's WPR is positively associated with proximity to the Metro while men's response is opposite (or ambiguous), if the reduction of commuting cost by the metro is much larger for women than men.

Our theoretical model basically follows the modelling strategy by Ahlfeldt et al. (2015) and Monte, Redding, and Rossi-Hansberg (2018), who study the residence-commuting choice of households within an urban area. Like their approach, we model the heterogeneity of individual choice following Eaton and Kortum (2002), and introduce an Fréchet distributed idiosyncratic welfare shock across destination and employment status that gives a convenient functional form for the destination choice in the equilibrium.

## 6.1 The Spatial Commuting Model

Let us assume a square-shaped city which consists of $J$ equally sized zones. In each zone $j$, $M_j$ number of men and $F_j$ number of women live for all $j$ for $1, 2, ..., J$. They don't move to other zones in the city, for the sake of simplicity. Thus, the residential population in each zone is fixed in the model. We denote gender as $G$, which takes $F$ or $M$ in what follows. For each region and gender, a fixed reservation wage $r_j^G$ is assured for every residents if an individual does not receive labour wage. Each person can work anywhere in the city and will be compensated with a zone $j$ specific wage $w_j$. If a resident in $j$ works in $j'$, she or he has to incur an iceberg type commuting cost $\tau^G_{jj'} \geq 1$.[13] Therefore, effective wage that a zone $j$ resident of gender $G$ working in $j'$ receives $w_{j'}/\tau^G_{jj'}$.

An individual consumes a homogeneous variety (numeràire) at an unity price. In addition, he or she draws an idiosyncratic utility shock for all the potential employment status, denoted by $\epsilon_{jm}(i)$. This is a shock of individual $i$ living in $j$, whose employment status is $m$. The employment status $m$ takes $H$ if $i$ chooses not to work and $m$ takes $j'$ if she commutes to $j'$ for work. $\epsilon_{jm}(i)$ follows Fréchet distribution with mean $B_m$, which corresponds to the average amenity level of $m$, and dispersion parameter $\eta$ with the CDF given by $F(\epsilon_{jm}) = \exp[-B_m\epsilon^{-\eta}]$. Given these idiosyncratic shocks ($\epsilon_{jm}$), the effective wage rates in all destination zones ($w'_j/\tau^G_{jj'}$), and the reservation wage level in the own residential zone ($r^G_j$), the individual $i$ chooses whether he or she works and where to commute so that his or her welfare is maximized. Thus the welfare of individual $i$ can be defined as

$$V_{G,j}(i) = \max\{\epsilon_{jH}r^G_j, \epsilon_{j1}\frac{w_1}{\tau^G_{j1}}, ..., \epsilon_{jJ}\frac{w_J}{\tau^G_{jJ}}\}. \tag{4}$$

Each zone produces a homogeneous product using labor and land. Land is a fixed endowment to zones. Let $N_j$ denote the supply of labor to city $j$ and $D_j$ is the land endowment. We assume a simple Cobb-Douglass production function;

$$Y_j = A_j N_j^{\beta} D_j^{1-\beta}. \tag{5}$$

where $A_j$ is $j$'s productivity shifter and $\beta \in (0, 1)$. The goods market is perfectly competitive, and workers regardless of their gender receives wage, $w_j$, which is equal to the

13. Within the same zone, we do not assume no commuting cost, meaning that $\tau^G_{jj} = 1$.

marginal productivity of labor.

In the equilibrium, the residents choose whether to work ($m = H$ if they do not work) and the destination of commuting if they work ($m = j'$). Labor and land are fully employed in each zone in the city, and the goods market clears. Therefore, the equilibrium of the urban economy can be defined by the equations below. Firstly, thanks to the property of Fréchet distribution of individual's idiosyncratic utility shock $\epsilon$, the probability that a gender $G$ resident in $j$ decides to commute to zone $j'$ is equivalent to the share of gender $G$ residents in $j$ commuting to $j'$ as follows

$$\pi_{jj'}^G = \frac{B_{j'}(w_{j'}/\tau_{jj'}^G)^\eta}{B_H(r_j^G)^\eta + \sum_k B_k(w_k/\tau_{jk}^G)^\eta}, \quad \forall G = \{F, M\}. \tag{6}$$

Given the above shares and the fixed population of men and women in each zone, the total labour supply to zone $j$ is then given by

$$N_j = \sum_{j'} \pi_{j'j}^M M_{j'} + \sum_{j'} \pi_{j'j}^M F_{j'}. \tag{7}$$

Finally, the equilibrium wage is equal to the marginal product of labour,

$$w_j = \beta A_j D_j^{1-\beta} N_j^{\beta-1}. \tag{8}$$

The equilibrium work participation rate (WPR) of gender $G$ in $j$ is 1 minus the inactive rate. From equation (6), this is given as

$$WPR_{G,j} = 1 - \frac{B_H(r_j^G)^\eta}{B_H(r_j^G)^\eta + \sum_k B_k(w_k/\tau_{jk}^G)^\eta}, \quad \forall G = \{F, M\}. \tag{9}$$

## 6.2 Comparative Statics of Transportation on WPR by Gender

Our main concern is how a heterogeneous change in commuting cost across gender ($t^G_{jj'}$) induced by the development of urban transportation network would affect the work participation rates of men and women in each zone ($WPR_{G,j}$). Especially, we are interested in whether there are cases where the decline in commuting cost has opposite WPR results for men and women.

Since the model is not analytically solvable, we conduct numerical simulations to examine its properties. We consider a square-shaped model economy consisted with total $J = \tilde{J} \times \tilde{J}$ tiles. For simplicity, we assume that the productivity ($A_j$), land ($D_j$), and average amenity level ($B_j$) take the value of 1 for every zone. Each zone is populated with a normalised population of men and women, namely $M_j = 1$ and $F_j = 1$, $\forall j = 1, ..., J$. There are two universal parameters in the model, the share of labour in production, $\beta$, and the shape parameter for the Fréchet distribution of destination preference, $\eta$. We set $\beta = 0.8$ and $\eta = 4$ following the literature.[14]

The transportation network is defined as the set of gender differentiated iceberg commuting cost between any pair of adjacent zones. For any pair of adjacent zones $j$ and $j'$, the link (node) between the two zones has either "traditional" or "metro" transit mode, denoted by $t(jj')$ which takes value 0 if the link $jj'$ has "traditional" transit or 1 if it has "metro" transit. Let $p^G_t$ denote a per unit distance traveling cost for a particular gender $G$ by a specific mode of transportation $t$. Traditional mode of transportation incurs $p^G_0 > 1$ of wage per unit distance for a gender $G$ commuter. Here, we are trying to replicate the situation before the Delhi Metro is commissioned. With the "metro", travelling one unit

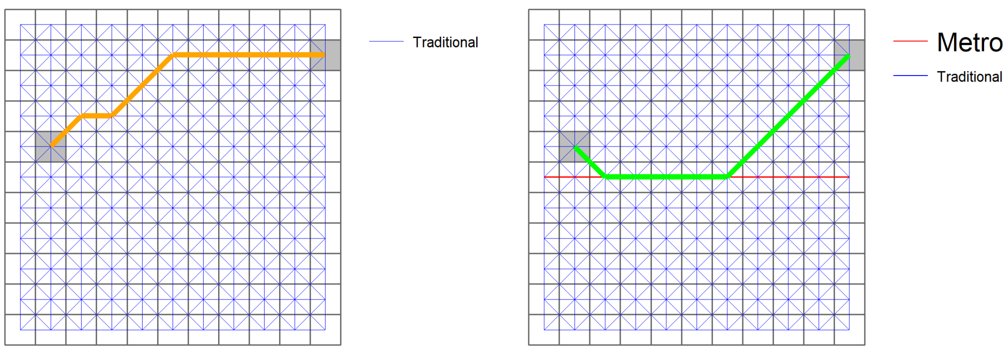
14. $\beta = 0.8$ refers to the choice by Ahlfeldt et al. (2015) in their calibration of the model which has a similar production assumption. Both Ahlfeldt et al. (2015) and Monte, Redding, and Rossi-Hansberg (2018) estimate the parameters corresponding to our $\eta$ for Berlin and the U.S. cities, respectively, and obtain the values between 3 to 5. Thus, we pick 4 for our analysis. Note that their shape parameters for the shock govern people's simultaneous choice of residence and commuting. Instead, in our model, residence is fixed.

distance costs $p_1^G > 1$, while we assume that the metro is cheaper in terms of the welfare cost of travelling than the traditional mode and that means $p_0^G > p_1^G$. The reduction of welfare cost is trying to represent an improved safety due to the Metro. In general, we denote the gender $G$ commuting cost between these two adjacent zones $j$ and $j'$ as $\tau_{jj'}^G = p_{t(jj')}^G d_{jj'}$, where $t(jj') = \{0, 1\}$ and $d_{jj'}$ is the distance between $j$ and $j'$. The commuting cost between the non-adjacent pairs of zones is defined as the least cost path to reach from $j$ to $j'$.

Figure 6 schematically depicts the city zones, its transport network, and the commuter's routing for the case of $\tilde{J} = 11$. The square tiles with black border lines are the zones. The centroids of adjacent zones are connected with blue lines that represent the nodes of the transportation network. For adjacency, we adopt the "queen" adjacency criteria which admits the two zones are adjacent even if only the corner is shared. Therefore, lines of diagonal directions are also included in the network. In panel (6a), we assume that the entire transportation network is served by the “traditional" mode and thus the entire network is colored in blue. It incurs $p_0 = 2$ of commuting cost per a unit distance. Let us consider the case of commuter's routing between the two grey colored zones. Under this environment without “metro", (an example of) the commuter's routing becomes the thick orange line, which realises the least cost. Instead, in panel (6b), we introduce a East-West “metro" line depicted in red. In this example, the unit distance cost is reduced to $p_1 = 1.5$ only on this red line. Travelling along the red line incurs fewer costs for commuters than passing through the blue traditional nodes. This will divert the least cost path between the two zones from that in the panel (6a) to the one like the green thick line in the panel (6b).

In the simulation analysis, we differentiate the traditional commuting cost of females and males. Reflecting the anecdotal context, we assume that the traditional welfare cost of

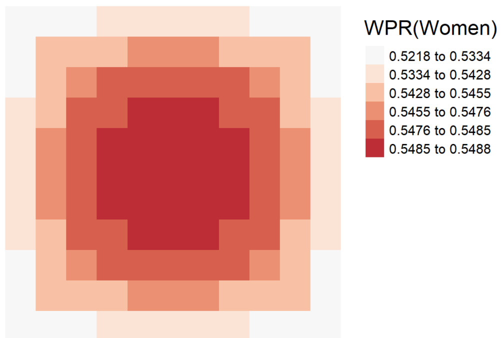
**Figure 6:** City Zones, Transport Network, and Commuters' Routing



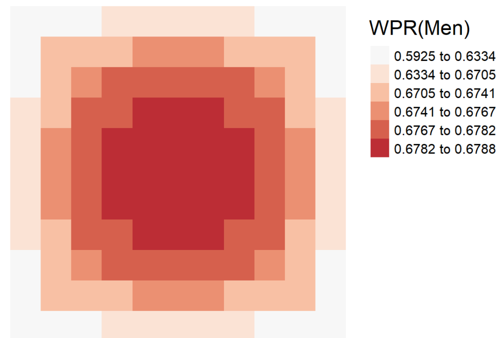

(a) Route without Metro (b) Route with Metro

*Source:* Authors
*Note:* In the panel (a), all the nodes on the network (in blue) requires $p_0 = 2$ commuting cost to travel. One of the least cost path between the two grey shaded zones is depicted as the orange thick line. Instead, in the panel (b), the travel cost on the East-West corridor in the middle which is depicted as a red line reduces to $p_1 = 1.5$, while the remaining nodes in blue stay at the level of $p_0$. The resulting least cost paths to travel between the two grey shaded zones changes to the green thick line.

**Figure 7:** WPR of Females and Males in the Initial Equilibrium (No Metro)




(a) Women's Initial WPR ($p_0^F = 2$)




(b) Men's Initial WPR ($p_0^M = 1.5$)

*Source:* Authors

*Note:* Spatial distributions of women's and men's WPR across zones, assuming that the entire urban transport network is traditional. To express women's disadvantage in the mobility in the initial equilibrium, commuting cost per unit distance is 2 for women and 1.5 for men on every node of the network.

commuting is higher for females than that for males. Specifically, we set $p_0^F = 2$ for females, while $p_0^M = 1.5$ for males. Figure 7 show the model's equilibrium WPRs for females and males given by equation (9), when the entire urban transport network is traditional as shown in Figure 6a. For both women and men, the WPR is higher in the central zones in the city, while it gradually reduces in the peripheral zones. Female WPR is much lower than for males. For females, the WPR ranges from 0.5218 to 0.5488. For males, the range shifts up to 0.5925 to 0.6788. These figures imply two things. Even on a featureless plain with equally distributed population and a featureless transport network, the residents of central location are more likely to work than those living in the periphery. In general, the higher the commuting cost is, the lower is the work participation rate. These results partially explain the situation of Delhi in 2001 that is depicted in Figure 3.

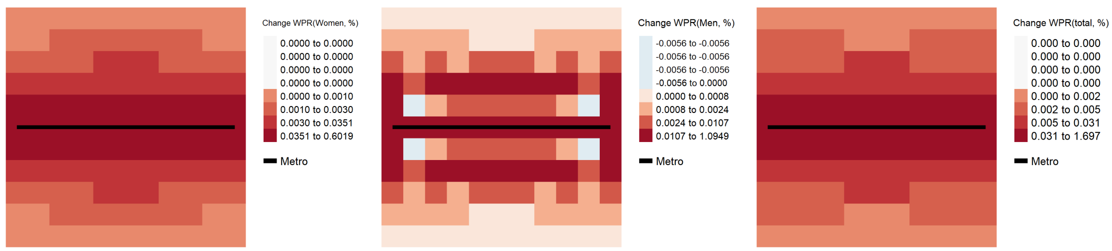
Figures 8, 9, and 10 depict the simulated changes in female and male WPR in response to the introduction of the Metro from East to West in the middle of the city, just as the red line in Figure 6b indicates. Figure 8 illustrates the impact of the Metro development when it reduces the commuting cost along the Metro corridor by 5% for both females and males. Namely, women's commuting cost reduces to 1.9 on the corridor (along the thick black line), while for men it becomes 1.425. For females as in the panel (a), the effect on the WPRs ranges from 0.0004 to 0.6019 percent, which are all positive but very marginal. And a clear "distance decay" pattern can be observed. Increasing the magnitude of commuting cost reduction for women will change not only the female WPR but also the male one. The male WPR shown in panel (b) responds in a more complicated way. While the positive impact is large in the immediate neighbourhood of the Metro alignment, the WPR interestingly reduces in a few spots locating relatively close to the Metro. In the majority of zones, the increase in the female WPR surpasses that of males. For the same magnitude of benefit (5 percent reduction in commuting cost), the group with severer initial deprivation will on average achieve larger gains. As in panel (c), the aggregate impact on the zonal employment is everywhere positive. In this case, the negative impact on male WPR in some zones is perfectly offset by the positive impact on female WPR.

In Figure 9 and Figure 10, we compare the cases of the Metro development that reduces female commuting costs more than that of males. Figure 9 illustrates the response of the WPRs when the Metro reduces female commuting cost along the Metro alignment to 1.425, which is the same level as the male metro commuting cost. With this relatively huge decline in female commuting cost on the Metro (28.75% reduction from the original), female WPR and male WPR show different responses. We observe an overall increase of female WPR, while male WPR declines in a large number of zones except for some specific places. Female WPR increases more in the central area where metro development happens. On the other hand, male WPR exhibits a more complicated response. Even in locations that are close enough to the Metro line, male WPR can decline (blue to grey shades in panel (b)). This shows a complex interdependent mechanism that the Metro development may deliver for women and men. For males in the immediate neighbourhood of the Metro alignment, the greater convenience for commuting leads more of them to seek employment. However, males in zones which are the second closest to the Metro line are crowded out from work due to the increased work participation of female residents and incoming commuters. In this case, the relationship between the distance to Metro and male work participation rates becomes ambiguous. Interestingly, the decline in male work participation slightly exceeds the female increase in a few locations. This is shown in panel (c) of Figure 9. Four zones near the both ends of the metro line exhibit an aggregate decline in overall WPR.

Figure 10 is a far more stringent case where the Metro serves much better for women than men and the female commuting cost reduces to 1.2 against the male cost of 1.425. In this case, while the results for females are qualitatively the same as that of Figure 9, the male WPR reduces almost everywhere in the city. The crowd out in the labor market contributes to reduce male work participation, especially within the central zones that the Metro most serves. The area with total employment decline expands in this case compared to the case in Figure 9, as shown in panel (c).

In summary, our empirical results can be at least partially explained by the mechanisms of this model - the gender differentiated commuting cost and interdependent relationship through the labor market. If the commuting cost reduction by the Metro is larger for females (who would be more constrained for mobility without it) than males, the adjustment through the local labor market will results in a positive effect of proximity to the Metro station on female WPR and an ambiguous effect on male WPR. Of course, we do not argue that the model describes the decisive mechanism that delivers our empirical results. There

**Figure 8:** The Change of WPR after the Commission of an East-West Metro: Case with 5 percent Reduction for both Females and Males on the East-West Corridor



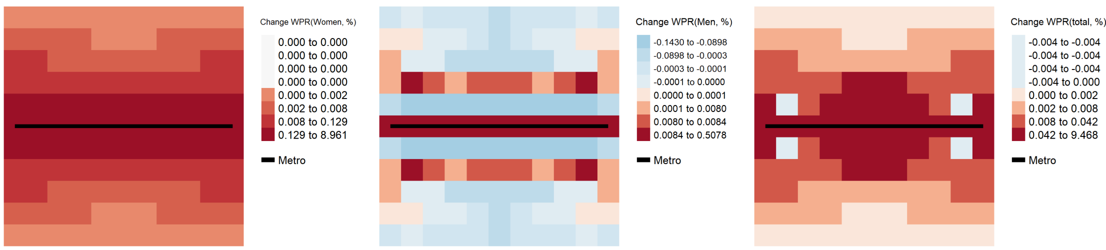

(a) Women's WPR Change (b) Men's WPR Change (c) Total WPR Change

*Source:* Authors

*Note:* Panels depicts changes in female and male WPR when the commuting costs of the East-West corridor reduces to $p_1^F = 1.9$ (by 5%) and $p_1^M = 1.425$ (by 5%) reduction, from the initial $p_0^G$ by the introduction of the Metro. Panel (a) is for females, panel (b) for males, and panel (c) for the total (females plus males), respectively.

could be a number of other mechanisms that are consistent with these results. This simulated model is the display of one possible mechanism. Especially, in the current analysis our model rules out endogenous residential choice of agent within the city. If people move in the city to maximize their utility, the effect of metro on female and male WPR can either be mitigated or amplified. Furthermore, we assume a single employment sector where both female and male workers compete. We can instead introduce multiple employment sectors so that gender sorting of working sector can be observed. In such a model, the key mechanism of our current results, the crowding out of male workers by the influx of female workers in the local labor market, may not happen.

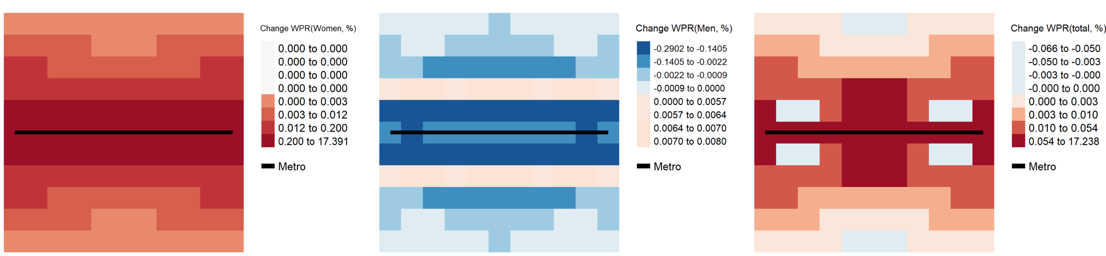
**Figure 9:** The Change of WPR after the Commission of an East-West Metro: Case where females and males achieve the same commuting cost on the Metro




(a) Women's WPR Change (b) Men's WPR Change (c) Total WPR Change

*Source:* Authors
*Note:* Panels depicts change in female and male WPR when the commuting costs on the East-West corridor reduces to $p_1^F = 1.425$ (by 28.75%) and $p_1^M = 1.425$ (by 5%) reduction, from the initial $p_0^G$. Panel (a) is for females, panel (b) for males, and panel (c) for the total (females plus males), respectively.

**Figure 10:** The Change of WPR after the Commission of an East-West Metro: Case where female commuting costs on the Metro become cheaper than male commuting costs




(a) Women's WPR Change (b) Men's WPR Change (c) Total WPR Change

*Source:* Authors
*Note:* Panels depicts change in female and male WPR when the commuting costs on the East-West corridor reduces to $p_1^F = 1.2$ (by 40%) and $p_1^M = 1.425$ (by 5%) reduction, from the initial $p_0^G$. Panel (a) is for females, panel (b) for males, and panel (c) for the total (females plus males), respectively.

## 7. Conclusion

In this study, we analyze the effect of the proximity to the Delhi Metro station which have opened up during the Phase I and Phase II of the project, from 2002 to 2011, on the work participation rate of females and males, using the Indian census that provides various demographic information of more than 300 geographical zones within Delhi. Thanks to the data structure with two pre-treatment period observations, we employ the Difference-in-Differences estimation controlling for a zone fixed effect, and jointly verify the common trend assumption during the pre-treatment periods. The overall results suggest that the proximity to the Metro station is positively related to female work participation, while the relationship between proximity to the Metro and male WPR is ambiguous, possibly having an opposite relationship to the case of females.

These findings provide suggestive evidence that the Metro encouraged females to participate in economic activities more than it did for males. This could be realized by, according to our conjecture, that the Delhi Metro might provide a safer mode of transportation that would benefit females who have suffered from safety problems more than males. With an parsimonious spatial urban model with commuting choice, we show that the larger reduction of commuting cost for females than males along with the Metro alignment can deliver important spatial patterns of this change in female and male WPRs that are similar to the ones empirically quantified.

However, we still need further investigation to know the causal link and the precise mechanisms behind it. More specifically, with the current dataset we cannot tell exactly why the positive effect on female rather than male economic participation is observed. At this stage, we only succeed in documenting the gender-heterogeneous correlation between proximity to the Metro and employment outcome. It is unclear whether the improved safety of commuting path has encouraged women to take a job outside of their home, since we do not directly observe their commuting choices. Alternative stories driven by labor demand can generate the same pattern of work participation rate. For instance, the Delhi Metro could have stimulated commercial activities around the Metro stations, such as retail shops, restaurants, offices, and so on. If some female oriented services (either by gender-wage gap or stakeholders' preference/discrimination) flourish in areas near stations, this would create more female employment opportunities than those for males. In this case, it would not be the safety of the Metro facility itself but the type of industries attracted to the premises of the Metro stations that would generate the observed pattern of female and male work participation rates. We leave remaining questions for future research with more detailed data.

## References

African Development Bank Group. 2009. *Checklist for Gender Mainstreaming in the Infrastructure Sector.* Nairobi: African Development Bank.

Ahlfeldt, Gabriel M., Daniel M. Sturm, Stephen J. Redding, and Nikolaus Worf. 2015. “The Economics of Density: Evidence from the Berlin Wall.” *Econometrica* 83 (6): 2127–89.

Andres, Luis A., Basab Dasgupta, George Joseph, Vinoj Abraham, and Maria Correia. 2017. “Precarious Drop: Reassessing Patterns of Female Labor Force Participation in India.” *World Bank Policy Research Working Paper* 8024:Washington DC: World Bank.

Angrist, J. D., and J. S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion.*

Asian Development Bank. 2013. *Gender Tool Kit : Transport. Maximizing the Benefits of Improved Mobility for All.* 95. Manilla: Asian Development Bank.

Autor, David H. 2003. “Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing.” *Journal of Laobr Economics* 21 (1).

Black, Dan A., Natalia Kolesnikova, and Lowell J. Taylor. 2014. “Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities.” *Journal of Urban Economics* 79:59–71.

Borker, Girija. 2017. “Safety First: Perceived Risk of Street Harassment and Educational Choices of Women.” PhD Thesis.

Eaton, Jonathan, and Samuel Kortum. 2002. “Technology, Geography, and Trade.” *Econometrica* 70 (5): 1741–79.

Gaduh, Arya, Tadeja Gracner, and Alexander D. Rothenberg. 2018. “Improving Mobility in Developing Country Cities: Evaluating Bus Rapid Transit and Other Policies in Jakarta.” mimeo.

Gimenez-nadal, J Ignacio, and Jose Alberto Molina. 2014. “Commuting Time and Labour Supply in the Netherlands A Time Use Study.” *Journal of Transport Economics and Policy* 48 (3): 409–26.

Gimenez-Nadal, J. Ignacio, and José Alberto Molina. 2016. “Commuting Time and Household Responsibilities: Evidence Using Propensity Score Matching.” *Journal of Regional Science* 56 (2): 332–59.

Glick, Peter. 1999. “Simultaneous Determination of Home Work and Market Work of Women in Urban West Africa.” *Oxford Bulletin of Economics and Statistics* 1.

Goel, Rahul, and Geetam Tiwari. 2016. “Access-egress and other travel characteristics of metro users in Delhi and its satellite cities.” *IATSS Research* 39 (2): 164–72.

Gutiérrez-i-Puigarnau, Eva, and Jos N. van Ommeren. 2010. “Labour supply and commuting.” *Journal of Urban Economics* 68 (1): 82–89.

Hyodo, Tetsuro, Cresencio Jr. M.Montalbo, Akimasa Fujiwara, and Sutanto Soehodho. 2005. "Urban Travel Behavior Characteristics of 13 Cities." *Journal of the Eastern Asia Society for Transportation Studies* 6:23–38.

Jaeger, David. A., Theodore J. Joyce, and Robert Kaestner. 2018. "A Cautionary Tale of Evaluating Identifying Assumptions: Did Reality TV Really Cause a Decline in Teenage Childbearing?" *Journal of Business & Economic Statistics* 21 (July): 1–10.

Jogori and UN Women. 2011. *Safe Cities Free of Violence Against Women and Girls Initiative: Report of the Baseline Survey Delhi 2010.* Technical report.

Kahn-Lang, Ariella, and Kevin Lang. 2019. "The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications." *Journal of Business and Economic Statistics:* 1–26.

Kawabata, Mizuki, and Yukiko Abe. 2018. "Intra-metropolitan spatial patterns of female labor force participation and commute times in Tokyo." *Regional Science and Urban Economics* 68 (November 2017): 291–303.

Kearney, Melissa S., and Phillip B. Levine. 2015. "Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing." *American Economic Review* 105 (12): 3597–3632.

Klasen, Stephan, and Janneke Pieters. 2015. "What explains the stagnation of female labor force participation in Urban India?" *World Bank Economic Review* 29 (3): 449–78.

Majid, Hadia, Ammar Malik, Kate Vyborny, Rania Nasir, Chuhang Yin, Mahrukh Khan, Fatima Khan, et al. 2018. "Infrastructure investments and public transport use Evidence from Lahore, Pakistan Infrastructure investments and public transport use: Evidence from Lahore, Pakistan," no. March.

Martínez, Daniel, Oscar A. Mitnik, Edgar Salgado, Lynn Scholl, and Patricia Yáñez-Pagans. 2018. "Connecting to Economic Opportunity: The Role of Public Transport in Promoting Women's Employment in Lima." *IZA Discussion Paper Series* December (12020): 44.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, migration, and local employment elasticities." *American Economic Review* 108 (12): 3855–90.

Onishi, Yumiko. 2017. *Breaking Ground: A narrative on the making of Delhi Metro.* 1–4. Japan International Cooperation Agency.

Peters, Deike. 2013. "Gender and Sustainable Urban Mobility." *Thematic study prepared for Global Report on Human Settlement:* Nairobi: UN–HABITAT.

Redding, S. J., and M. A. Turner. 2015. "Transportation Costs and the Spatial Organization of Economic Activity." *Handbook of Regional and Urban Economics* 5 (1): 1339–98.

Safetipin. 2016. *Using Data to Build Safer Cities.* Technical report.

Takaki, Keiichi, and Yoshimi Hayashi. 2012. *India Ex-Post Evaluation of Japanese ODA Loan "Delhi Mass Rapid Transport System (I)-(VI)".* Technical report I.

UN Women. 2014. *Ensuring Safe Public Transport With and for Women and Girls in Port Moresby Papua New Guinea.* 42. Papua New Guinea: UN Women.

United Natations. 2019. *World Urbanization Prospects: The 2018 Reivion.* Technical report. New York: United Nations, Department of Economic and Social Affairs, Population Division.

Uteng, Tanu Priya. 2011. "Gender and Mobility in the Developing World." *World Development Report Background Paper.*

World Bank. 2010. "Making Infrastructure Work for Women and Men_a Review of World Bank Infrastructure Projects (1995-2009)."

Zax, Jeffrey S. 1991. "Compensation for commutes in labor and housing markets." *Journal of Urban Economics* 30 (2): 192–207.

# A. Additional Information for Descriptive Statistics

**Table A.1:** statistical test for the difference of distribution

| | (1)<br>female<br>(91 vs 01) | (2)<br>female<br>(01 vs 11) | (3)<br>male<br>(91 vs 01) | (4)<br>male<br>(01 vs 11) |
|---|---|---|---|---|
| | *p*-values | | | |
| T-test (mean) | 0.000 | 0.0013 | 0.000 | 0.0056 |
| K-S test | 0.000 | 0.007 | 0.000 | 0.120 |

*Note:* "T-test" reports the p-values the T-test to compare means of the WPR across years. "K-S test" reports the results (p-values) of the Komogorov-Smirnov test for comparing two distributions of WPR across years. Column (1) is for female WPR between 1991 and 2001. Column (2) is for male WPR between 2001 and 2010. Column (3) and (4) report the same for men.

# B. Estimates without Controls

This section provides the estimation results without control variables as a robustness check to our main results in Section 5. If our control variables (household size, children share, female to male literacy ratio, and female to male ratio of scheduled caste) are also the outcome of the development of the Delhi Metro, this may bias the estimates for our variable of interest. We only keep our variables of interest, the distance to the Metro station (2011) and its lead term, then perform a fixed effect estimation. Table B.2 and Table B.3 show the estimation results for female and male WPR, respectively.

**Table B.2:** Female WPR, without Controls

| VARIABLES | (1) OLS ALL | (2) OLS d < 10km | (3) OLS d < 5km | (4) OLS d < 10km & CBD > 10km | (5) OLS d < 5km & CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | -0.00782*** | -0.00945*** | -0.00617*** | -0.0107*** | -0.00642*** |
| | (0.00149) | (0.00167) | (0.00149) | (0.00236) | (0.00205) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.00204 | -0.00155 | 0.000932 | -0.00115 | 0.00312 |
| | (0.00234) | (0.00284) | (0.00302) | (0.00368) | (0.00409) |
| Constant | 0.0708*** | 0.0725*** | 0.0738*** | 0.0661*** | 0.0663*** |
| | (0.00123) | (0.00128) | (0.00112) | (0.00179) | (0.00164) |
| Observations | 1,016 | 957 | 808 | 663 | 514 |
| R-squared | 0.157 | 0.151 | 0.220 | 0.101 | 0.152 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Year Dummy | YES | YES | YES | YES | YES |
| Adj-R | 0.153 | 0.147 | 0.217 | 0.0957 | 0.146 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

**Table B.3:** Male WPR, without Controls

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | OLS | OLS | OLS | OLS | OLS |
| | ALL | d < 10km | d < 5km | d < 10km | d < 5km |
| VARIABLES | | | | & CBD > 10km | & CBD > 10km |
| Dist. to Metro(2011, $\beta$) | 0.00697*** | 0.00839*** | 0.00727*** | 0.0143*** | 0.0135*** |
| | (0.00243) | (0.00234) | (0.00274) | (0.00332) | (0.00400) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.0273*** | -0.0285*** | -0.0315*** | -0.0190*** | -0.0232*** |
| | (0.00451) | (0.00480) | (0.00628) | (0.00650) | (0.00843) |
| Constant | 0.439*** | 0.447*** | 0.465*** | 0.424*** | 0.446*** |
| | (0.00197) | (0.00197) | (0.00207) | (0.00261) | (0.00291) |
| Observations | 1,016 | 957 | 808 | 663 | 514 |
| R-squared | 0.248 | 0.251 | 0.218 | 0.289 | 0.259 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Year Dummy | YES | YES | YES | YES | YES |
| Adj-R | 0.245 | 0.248 | 0.214 | 0.285 | 0.254 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

**Table B.4:** WPR gap (female to male), without Controls

| VARIABLES | (1)<br>OLS<br>ALL | (2)<br>OLS<br>d < 10km | (3)<br>OLS<br>d < 5km | (4)<br>OLS<br>d < 10km<br>& CBD > 10km | (5)<br>OLS<br>d < 5km<br>& CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | -0.0207*** | -0.0257*** | -0.0151*** | -0.0330*** | -0.0193*** |
| | (0.00449) | (0.00491) | (0.00396) | (0.00692) | (0.00550) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.00882 | -0.00738 | 0.00342 | -0.00972 | 0.00685 |
| | (0.00573) | (0.00666) | (0.00651) | (0.00899) | (0.00914) |
| Constant | 0.161*** | 0.163*** | 0.157*** | 0.158*** | 0.148*** |
| | (0.00290) | (0.00295) | (0.00228) | (0.00415) | (0.00338) |
| Observations | 1,016 | 957 | 808 | 663 | 514 |
| R-squared | 0.177 | 0.177 | 0.256 | 0.132 | 0.185 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Year Dummy | YES | YES | YES | YES | YES |
| Adj-R | 0.174 | 0.173 | 0.253 | 0.127 | 0.179 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

**Table B.5:** WPR for total adult population, without Controls

| VARIABLES | (1)<br>OLS<br>ALL | (2)<br>OLS<br>d < 10km | (3)<br>OLS<br>d < 5km | (4)<br>OLS<br>d < 10km<br>& CBD > 10km | (5)<br>OLS<br>d < 5km<br>& CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | 0.00115 | 0.00104 | 0.00146 | 0.00431* | 0.00525* |
| | (0.00168) | (0.00169) | (0.00189) | (0.00242) | (0.00273) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.0170*** | -0.0173*** | -0.0180*** | -0.0139** | -0.0153** |
| | (0.00340) | (0.00374) | (0.00466) | (0.00606) | (0.00756) |
| Constant | 0.274*** | 0.279*** | 0.290*** | 0.265*** | 0.278*** |
| | (0.00148) | (0.00150) | (0.00155) | (0.00203) | (0.00222) |
| Observations | 1,016 | 957 | 808 | 663 | 514 |
| R-squared | 0.195 | 0.182 | 0.168 | 0.198 | 0.186 |
| Number of id | 342 | 322 | 271 | 224 | 173 |
| Year Dummy | YES | YES | YES | YES | YES |
| Adj-R | 0.192 | 0.179 | 0.163 | 0.193 | 0.180 |

*Note:* Standard errors are clustered at the individual zone
*** p<0.01, ** p<0.05, * p<0.1
"d < km" if sample zones with distance to Phase I - IV stations within x km
"CBD < km" if sample zones locate further than x km from the CBD

## C. Discussion on Data Interpolation

As argued in Section 3, the geographical boundaries of zones in Delhi have not stayed constant during the three rounds of the census, thus we have to interpolate the observed statistics in 2001 and 2011 so that the boundaries are consistent with those of 1991.

We therefore examine the sensitivity of our results to the interpolation method by comparing our main results in Section5 with the case where we limit the estimation sample only to the zones with consistent boundaries throughout 1991 to 2011. Out of 342 sample zones, 222 keep their boundaries constant across three periods. We repeat the same estimations for the WPR of females and males with this constant boundary subset. The results are shown in Table C.6 and Table C.7. The results generally support the prediction on the direction of bias.

Firstly, compared to the main estimates for female WPR shown in Table 2, the estimates with the boundary consistent subsets show the qualitatively similar results (Table C.6). For all the five specifications, the effect of the distance to metro station is negative, and the magnitude is about twice as large as that in Table 2. This implies that using the interpolated data gives to smaller estimates for the *positive* effect of the proximity to the Metro station, which is consistent with the explanation with the illustrated example in the Appendix C.

Table C.7 shows the results for male WPR with the same subset. While the coecients on the "Distance to Metro (2011)" are all positive for our main estimation, the results with the subset are neither positive nor significant. For males, the results are less stable across different specifications. The estimates for male WPR are sensitive to the sample choice as well as interpolation method for inconsistent zone boundaries.

**Table C.6:** Female WPR only with consistent boundary zones

| VARIABLES | (1)<br>OLS<br>ALL | (2)<br>OLS<br>d < 10km | (3)<br>OLS<br>d < 5km | (4)<br>OLS<br>d < 10km<br>& CBD > 10km | (5)<br>OLS<br>d < 5km<br>& CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | -0.0123*** | -0.0148*** | -0.0134*** | -0.0131*** | -0.0119*** |
| | (0.00212) | (0.00233) | (0.00255) | (0.00219) | (0.00223) |
| Dist. to Metro(1991, $\beta_{-1}$) | 0.000202 | 0.000573 | 0.00517 | -0.00377 | 0.000738 |
| | (0.00362) | (0.00416) | (0.00429) | (0.00347) | (0.00365) |
| Household Size | -0.0881** | -0.0892** | -0.0725* | -0.101** | -0.0808* |
| | (0.0393) | (0.0435) | (0.0411) | (0.0500) | (0.0478) |
| Children Share | -0.156*** | -0.166*** | -0.161*** | -0.153*** | -0.154*** |
| | (0.0252) | (0.0262) | (0.0295) | (0.0213) | (0.0232) |
| Female to male literacy ratio | -0.0966** | -0.0989** | -0.0672 | -0.124*** | -0.0946** |
| | (0.0397) | (0.0397) | (0.0424) | (0.0335) | (0.0379) |
| Female to male SC ratio | -0.0556** | -0.0674** | -0.0389 | -0.0800** | -0.0480 |
| | (0.0247) | (0.0314) | (0.0305) | (0.0309) | (0.0293) |
| Constant | -0.102 | -0.116 | -0.130 | -0.0786 | -0.109 |
| | (0.0881) | (0.0954) | (0.101) | (0.0932) | (0.0934) |
| Observations | 646 | 588 | 441 | 552 | 405 |
| R-squared | 0.417 | 0.428 | 0.388 | 0.524 | 0.464 |
| Number of id | 222 | 202 | 151 | 190 | 139 |
| Year Dummy | YES | YES | YES | YES | YES |
| Adj-R | 0.412 | 0.422 | 0.379 | 0.519 | 0.456 |

*Note:* Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
d< $x$km if sample zones with distance to Phase I - IV stations within x km
CBD< $x$ km if sample zones locate further than x km from the CBD

**Table C.7:** Male WPR only with consistent boundary zones

| VARIABLES | (1)<br>OLS<br>ALL | (2)<br>OLS<br>d < 10km | (3)<br>OLS<br>d < 5km | (4)<br>OLS<br>d < 10km<br>& CBD > 10km | (5)<br>OLS<br>d < 5km<br>& CBD > 10km |
|---|---|---|---|---|---|
| Dist. to Metro(2011, $\beta$) | -0.00394 | -0.00367 | -0.00603 | -0.00120 | -0.00416 |
| | (0.00313) | (0.00336) | (0.00394) | (0.00304) | (0.00300) |
| Dist. to Metro(1991, $\beta_{-1}$) | -0.00255 | -0.00479 | -0.00454 | -0.0109** | -0.0115*** |
| | (0.00583) | (0.00643) | (0.00728) | (0.00466) | (0.00426) |
| Household Size | -0.378*** | -0.360*** | -0.349*** | -0.374*** | -0.358*** |
| | (0.0424) | (0.0427) | (0.0410) | (0.0456) | (0.0440) |
| Children Share | -0.0709* | -0.0783** | -0.0847** | -0.0635* | -0.0830*** |
| | (0.0368) | (0.0378) | (0.0403) | (0.0323) | (0.0202) |
| Female to male literacy ratio | 0.0198 | 0.0116 | 0.0388 | -0.0337 | -0.0210 |
| | (0.0512) | (0.0507) | (0.0639) | (0.0208) | (0.0244) |
| Female to male SC ratio | 0.0543 | 0.0371 | 0.0590 | 0.0452 | 0.0702 |
| | (0.0504) | (0.0525) | (0.0570) | (0.0513) | (0.0526) |
| Constant | 0.909*** | 0.871*** | 0.864*** | 0.908*** | 0.866*** |
| | (0.115) | (0.118) | (0.119) | (0.109) | (0.0834) |
| Observations | 646 | 588 | 441 | 552 | 405 |
| R-squared | 0.504 | 0.503 | 0.512 | 0.610 | 0.622 |
| Number of id | 222 | 202 | 151 | 190 | 139 |
| Year Dummy | YES | YES | YES | YES | YES |
| Adj-R | 0.500 | 0.497 | 0.505 | 0.606 | 0.616 |

*Note:* Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
d< $x$km if sample zones with distance to Phase I - IV stations within x km
CBD< $x$ km if sample zones locate further than x km from the CBD

**Abstruct (in Japanese)**


要約

開発途上国では、伝統的に女性の移動が制約されていることがあるが、デリー・メトロはそうした開発途上国における都市の大量輸送インフラの代表例の一つである。本論文では、デリー・メトロが女性と男性の労働参加率に与える効果を、三期間（1991、2001、2011 年）のデリー市内の町レベルの区域パネルデータを用いて分析する。このデータでは個々の住民の特性を詳細に理解するには限界があるものの、区域ごとの固定効果を制御した上で差の差推定法を用い、平行トレンド仮定も検定する。推計結果としてはデリー・メトロの駅に近い区域ほど女性の労働参加率が有意に増加する一方、男性の労働参加率への影響は曖昧で、反対の符号をもつ可能性があることを示した。複数のメカニズムがこの結果を説明し得ると考えられるものの、理論的な都市通勤モデルを用いて、女性の通勤コストが男性に比して大幅に減少する場合（例えばより安全な通勤手段を提供する等）を検証したところ、実証的に定量化された傾向と同じ結果が得られることが確認された。概して、本論文の結果は都市交通インフラが包摂的な成長と貧困削減に与える定量分析の一連の文献に関連すると考える。



**キーワード：** インド、ジェンダーギャップ、均衡通勤モデル