Impact of the Proximity to the Delhi Metro on Work Participation of Female and Male*

Mai SEKI[†] and Eiji YAMADA[‡]

December 18, 2018

Abstract

In this paper, we analyze the impact of Delhi Metro on the work participation rate of females relative to males, to provide quantitative evidence on whether a high quality urban transportation contributes to reduce gender gap in economic participation. Using Primary Census Abstract (1991, 2001, and 2011) combined with map information of towns and metro alignments to construct accessibility measures, we examine whether the proximity to metro stations contributes to the area's growth in non-agricultural work participation for females in contrast to males. Our results indicate that the proximity to the Delhi Metro stations significantly increases the area's female work participation rate relative to male. Overall, our results hinge upon the literature on quantification of the contribution of urban transport infrastructure towards the inclusive growth and poverty reduction.

*We are thankful to the seminar participants at DIME workshop in Lisbon 2017, GRIPS 2017, Japan Evaluation Society Annual Conference 2017, ISEC Conference in Bangalore 2018, Prof. Takashi Kurosaki, Prof. Yasuyuki Sawada, Prof. Gilles Duranton, Prof. Kensuke Teshima, Prof. Naoya Sueishi, Dr. Yi Jiang, and Dr. Rana Hassan, for their helpful comments and suggestions. All errors are ours. This 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.

†Ritsumeikan University, maiseki@fc.ritsumei.ac.jp

‡JICA Research Institute, Yamada.Eiji@jica.go.jp

1 Introduction

In developing countries, an urbanization has progressed rapidly and more than the half of the world population already live in the urban areas as of 2014 (United Nations, 2014). To mitigate traffic congestions accompanied by the rapid urbanization, many countries are investing in urban transportation. While an overall mobility of residents improves and city's capacity continues to expand by those investments, gender inequality in public transportation has been remaining as a major issue (Peters, 2013; Uteng, 2011; Hyodo et al., 2005). According to these studies, females in urban areas of the developing countries go out of home less frequently, and depend more on public transportations than male counterparts. This indicates that a provision of safe and accessible public transportation could potentially improve female mobility, which leads to active participation of female to the economy.

In fact, a gender mainstreaming in infrastructure projects of developing countries has gained policy attentions over the past decade (Asian Development Bank, 2013; African Development Bank Group, 2009; UN Women, 2014; World Bank, 2010). However, there are still limited numbers of research quantifying the development impact of such policies. In the urban economics, there are studies discussing the gender heterogeneity in commuting time to work and its impact on labor supply (Gutiérrez-i Puigarnau and van Ommeren, 2010; Gimenez-nadal and Molina, 2014; Gimenez-nadal et al., 2015; Zax, 1991; Black et al., 2014; Kawabata and Abe, 2018); however, they do not necessarily focus on public transportations given country contexts, except the one by Abe and Kawabata analyzing the commuting and labor supply patterns of a married couples resident in Tokyo great metropolitan area. In the literature of transportation, urban planning or geography, there are studies documenting 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); however, they do not necessarily aim to address a causal relationship. Gaduh et al. (2018); on the other hand, estimate an equilibrium model of commuting choices with endogenous commuting times to assess the impact of counterfactual transportation policies, using the commuter information of Jakarta's Bus Rapid Transit (BRT) system. Their findings on gender-heterogeneous impact of the proximity to BRT on commuting time motivates our study to examine a heterogenous impact of the proximity to public transportation on labor supply by gender. In the literature of impact evaluation of the transportation, rural roads or intra-city highways and railways have been a focus of evaluation and very few impact evaluations on urban transportation exist (Seki, 2016). Among those, the most relevant analysis, which is ongoing, is the field-experiments 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.

In this paper, we analyze the impact of Delhi Metro on the work participation of female relative to male, to provide quantitative evidence on whether a high quality urban transportation contributes to improve female economic participation. Here, we focus on the Delhi Metro for three reasons. Firstly, Delhi is 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). In fact, Borker (2017) finds that safety of school-commuting route has direct impact on the university choice among the female students in the city of Delhi. Secondly, India faces a challenge for 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). Lastly, Delhi Metro can be one of the best candidates to analyze the impact of "high quality" urban transport infrastructure in developing countries, given its reputations for service standards, not only for its stability and convenience, but also for the safety and comfortability for female passengers. Based on the interviews of the users, the introduction of Delhi Metro is known that it drastically changed the transportation choice of women, due to the high standard of safety in the Metro system (Takaki and Hayashi, 2012; Onishi, 2017). Based on these reasons, we hypothesize the introduction of a safe mode of public transportation in Delhi would have had a non-negligible impact on female labor supply relative to the male, combined with other factors (i.e., residential relocations, family-level joint labor supply decision and/or compositional change in labor demand).

We use the Primary Census Abstract (PCA) which provides various tabulations from Population Census data at the fifth administrative level (town and village level). We construct a panel of PCA zones for three consecutive census years, 1991, 2001, and 2011. Furthermore, we calculate an accessibility measures from each PCA zone to the nearest metro station, using GIS information of PCA zones and the alignment of the Delhi Metro. With the calculated treatment variable, a proximity to the Delhi metro, we conduct a difference-in-differences (DID) analysis in order to assess whether the proximity to metro stations contributes to the area's growth of female participation to non-agricultural economic activities relative to males. This is an outcome measure which capture gender-heterogeneous impact from the Delhi Metro system. Our results indicate that the proximity to the Delhi Metro improve the economic participation of females more than that of males. While our data have some limitations for making rigorous causal inferences, especially for disentangling the mechanism behind the results, our study is the one of the first attempts to quantitatively measure the gendered implication of large scale urban transport development in the context of mega cities in developing countries.

The rest of 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. Section 6 discusses the limitation of our method and potential direction of future research.

2 Background of Delhi Metro

Among the rapidly urbanizing developing countries, India is expected to add 404 million urban population from 2014 to 2050, and its capital city, Delhi, is already the second largest city in the world, recording the population of 25 million. As the country's third urban mass rapid transit system (MRT), Delhi metro project has been developed over the past fifteen years. The first phase of Delhi Metro project consisted of 58 stations covering 65 km and commissioned during 2002-2006. Following the Phase I of the project, Phase II built 85 stations covering 125 km and commissioned during 2008-2011. Currently, Phase III project is under construction and it is expected to cover 106 km and Phase IV is under planning stage. Intuitively, the zones close to these Phase I and II metro stations are the "treatment" group in our analysis. Meanwhile, these "to-be-comissioned" Phase III and IV metro lines will be later utilized to refine our analysis, by restricting the "control group" zones. [1]

The novelty of Delhi Metro project is the fact that they had 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 to provide a safe public mass urban transportation for the citizens of Delhi. Overall, the Delhi Metro has gained the reputation for its high standard of facility and operation which ensures the safety and comfortability for female passengers (Takaki and Hayashi, 2012; Onishi, 2017).

[1] Alternative way of refining our analysis is to utilize "planned (but not constructed due to technical reasons uncorrelated with outcome variable)" metro lines. However, there was no major change in the plan for the case of Delhi Metro, so such approach is not taken.

Table 1: Results Summary of JICA's Beneficiary Survey (2016)

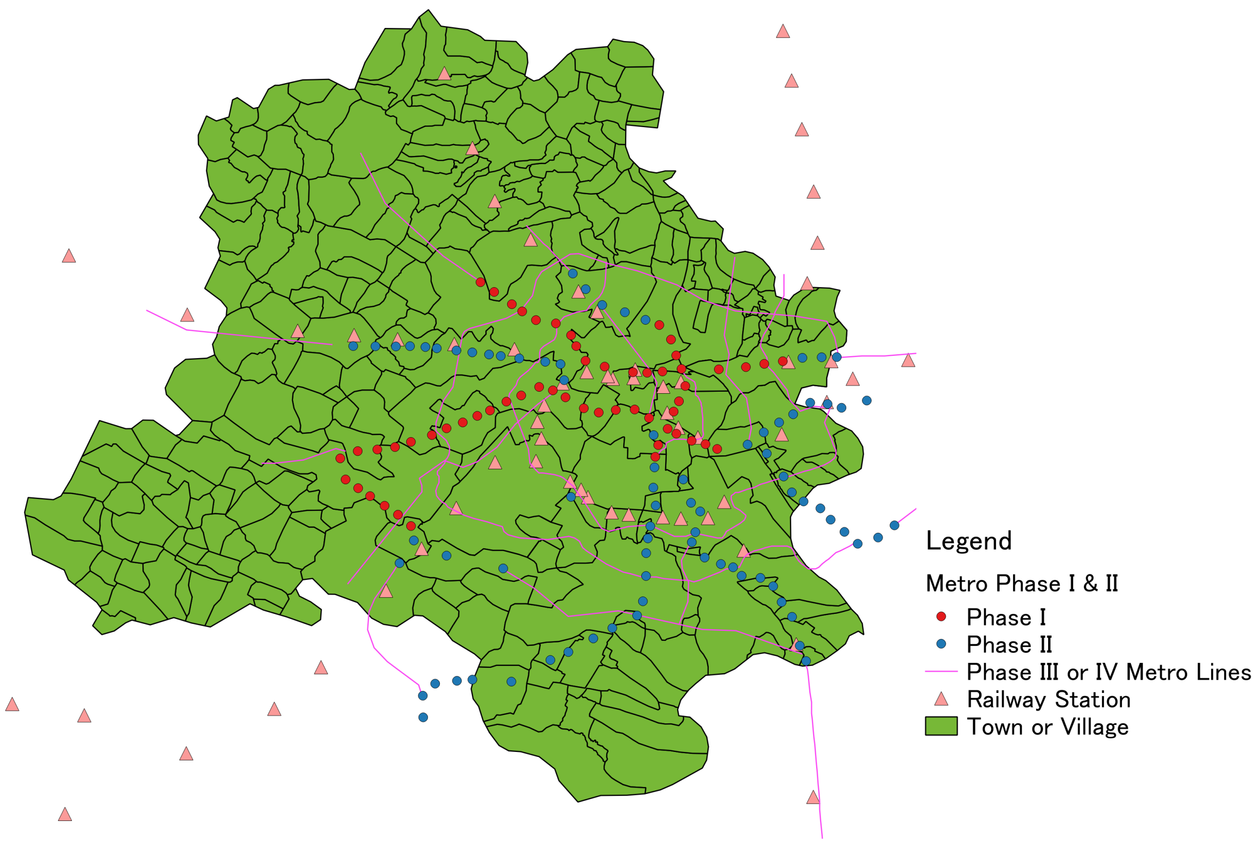
Do you think the Metro...?	Percentage of "Yes" Women ($N = 34$)	Men ($N = 116$)
helped women go out more frequently	100%	94 %
improved public security nearby station	88%	84 %
helped people go out after dark?	88%	79 %

Prior to the introduction of Delhi metro, female safety concerns in public transportation system had been severe 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 such as employment and schooling in optimal locations, or other activities such as household choirs, family visits, or leisure, it is not difficult to hypothesize that the limitation of safe modes of transportation was taxing for women to get access to certain social and economic opportunities. Given such a context, introduction of relatively safe public transportation system has had a potential impact to drastically change female behavior in Delhi.

This pro-female impact of the Delhi Metro has been anecdotally supported. For example, a beneficiary survey of the 150 Metro users and residents nearby stations, conducted by JICA in 2016, almost all of the respondents gave positive answers to the questions about the Metro's impact on public safety and opportunity for females, as summarised in the Table 1. In addition, several supporting comments from the respondents are reported, such as; "The Metro helps women go out alone", "The Metro saves time and is safe after dark. It is safer than the buses", and "Thanks to the Metro, my parents allowed me to start my job", etc.

In this paper, we focus on the first two Phases, I and II, to examine its impact on

Figure 1: Delhi Metro Stations



female work participation due to the timing of the data availability.

3 Data

We use the Primary Census Abstract (PCA) of India's Population Census, published by the Office of the Registrar General and Census Commissioner, Ministry of Home Affairs of 1991, 2001, and 2011. The PCA is an aggregate of population census enumeration at the level of a small local administrative unit, up to the fifth administrative level. 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.[2]

To represent economic participation of each gender 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)[3] divided by the adult population[4], for each gender. This indicator is different from labor force participation rate (LFPR). While the denominator

[2]We carry out the interpolation as follows. Suppose a zone i in 1991 boundary overlaps with n zones in 2001 boundary $j = 1, ..., n$. The area size of i is denoted by A_i and for j it is denoted by B_j. Let b_j represent the size of the area the zone j intersects with the zone i. Let b'_j the area size of the rest of the territory of j (i.e. which does not intersect with i). Then, by definition, $B_j = b_j + b'_j$ and $A_i = \sum_j^n b_j$. Now, suppose we want to interpolate a statistic x (e.g. population) in 2001 to be consistent with 1991 boundary. We calculate the interpolated value of statistic x for a 1991 geographical unit i, which is represented by $\tilde{x}_i$, by

$$\tilde{x}_i = \sum_j^n \frac{b_j}{B_j} x_j$$

[3]"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 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".

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

of LFPR is usually an working-age population above the age of 15, the denominator of WPR is (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. Other information used from the PCA tables is the number of household, total population, the number of children, the number of household, the number of literal residents, the number of residents scheduled caste (each by gender).

Our treatment variable is the proximity of a zone (town and villages based on 1991 administrative boundary) 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 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 $d_{k(i)}$. (iii) The average distance to the nearest Metro station(s) of the zone i, D_i, is then calculated as

$$D_i = \frac{\sum_{k(i)} d_{k(i)}}{N_i} \tag{1}$$

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 in early years during 2002-2011. Average distance measures to the railway stations already existed before 2001 and access to the Metro Phase III and IV (only under planning phase in 2011) are also calculated in the same manner to better define the comparison group which is more likely to share the similar unobserved characteristics regardless of the assigned treatment.

The descriptive statistics is shown in the Table 2. 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 is shorter with 3.3 km.

Female WPR has been substantially lower than that of male's throughout two decades since 1991. However, the average female WPR has increased from 5.3 % in 1991 to 7.9 % in 2011, while male's WPR has grown from 40.4% to 45.3% during the same period.

Figure 2 depicts kernel density estimates for the distribution of female and male WPR for years 2001 and 2011. First, we can observe that the WPR distributions are distinctly different across genders for both years (before and after the commission of Delhi metro Phase I and II). 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 WPR distribution towards right among females. This is suggesting that the rate was improved almost everywhere in the distribution for the females. As for males, we do not observe such one-directional change from 2001 to 2011.

Figure 3 shows the spatial distribution of WPR of female and male for 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 female WPR. The bottom two, 3c and 3d are those for male. All these four maps indicate high spatial and serial correlation of WPR.

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

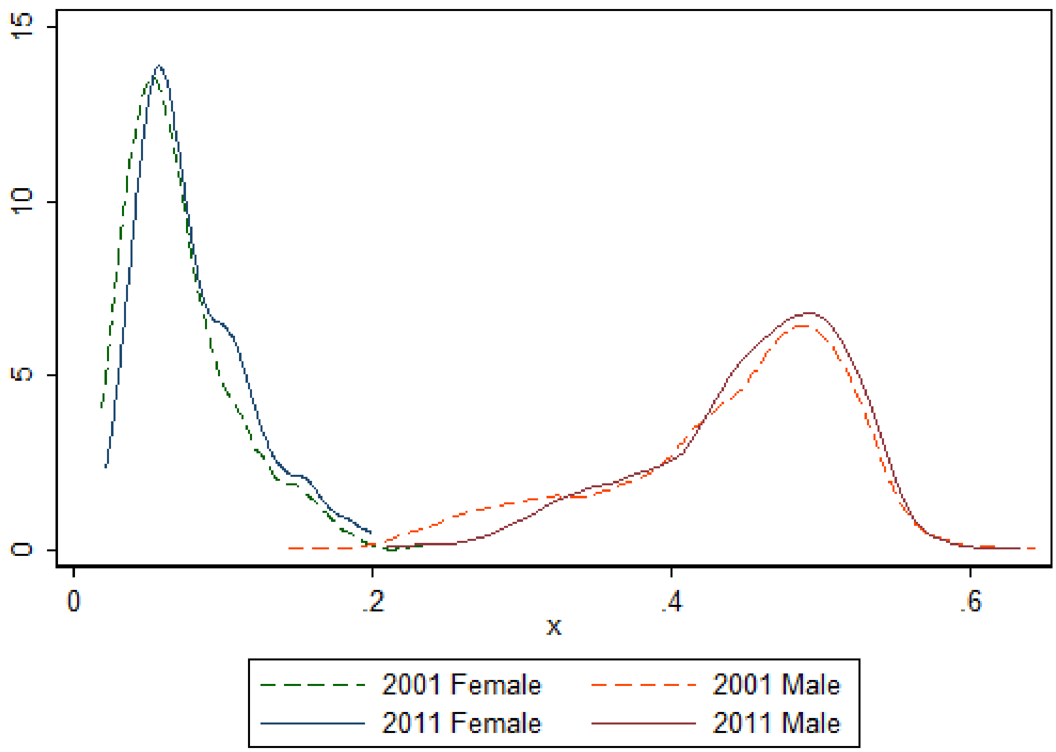
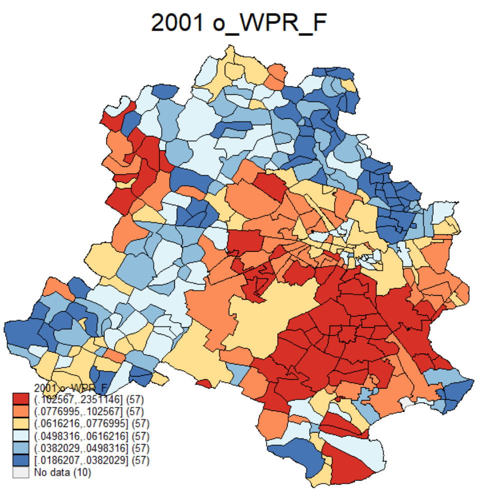


Table 2: (a) Summary Statistics: Level

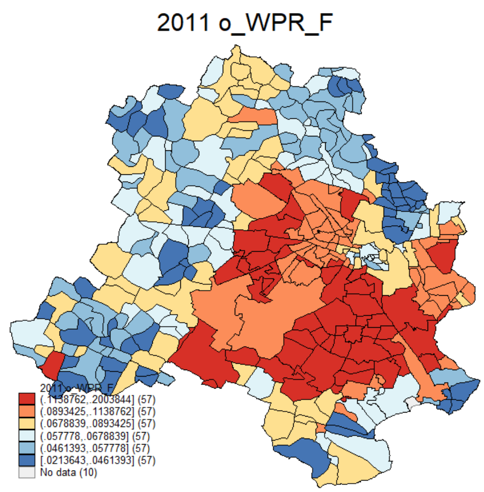
	1991			2001			2011		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	N	mean	sd	N	mean	sd	N	mean	sd
Dist. to Phae 1 or 2 Metro St.							342	5.239	4.763
Dist. to Phae 3 or 4 Metro St.							342	3.274	3.145
female WPR	332	0.0531	0.0599	342	0.0706	0.0369	342	0.0791	0.0371
male WPR	332	0.404	0.131	342	0.439	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

Figure 3: Spatial Distribution of WPR for female and male, in 2001 and 2011

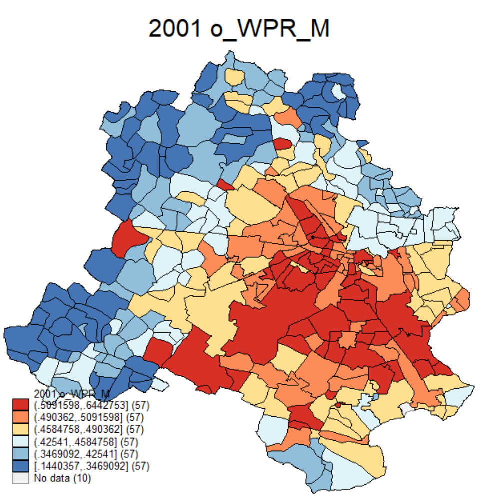
(a) 2001 Female WPR



(b) 2011 Female WPR



(c) 2001 Male WPR



(d) 2011 Male WPR

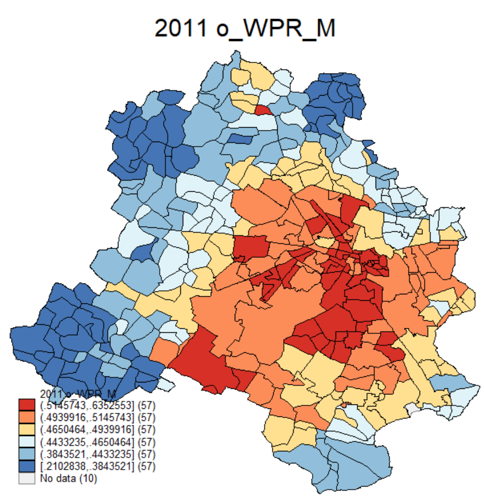
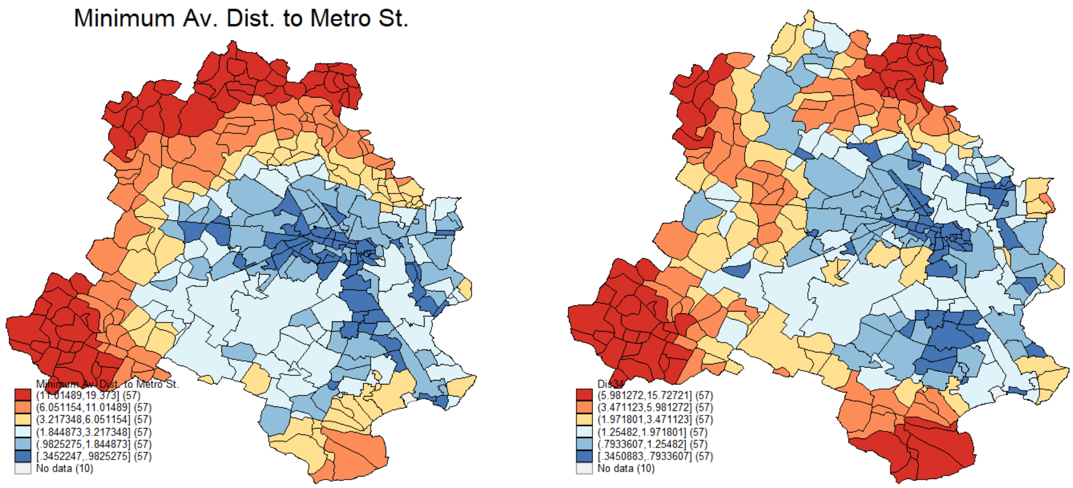


Figure 4: Distance to Commissioned and Planned Metro Stations

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



4 Empirical Strategy

A goal of this paper is to empirically test the anecdotes that the Metro facilitates the female participation of economic activities in urban Delhi. While we cannot separately identify the impact of different mechanisms, we aim to capture heterogeneous impact of metro by gender as the first step. More specifically, we investigate whether the zones closer to the Delhi Metro station has observed more increase in female work participation (than that of males). 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 ($= WPR(female)/WPR(male)$), or "WPR ratio" in short; and WPR for total workers (sum of female and male). These four outcome measures have different roles in interpretation. Female WPR and male WPR will tell us the impact on each gender separately, unconditional on the impact on the other gender. With the WPR ratio, the estimation result will capture the impact on females relative to males. The rationale of using this outcome measure is to capture the gender-heterogenous benefits of Delhi metro (e.g., safety from sexual violence).

Other than the impact through the safety feature of Delhi metro, there are a couple of other mechanisms that could generate gender-heterogeneous impacts. Firstly, labor demand might change by the introduction of the Metro and that could be gender-heterogenous. Secondly, a reduction of congestion and travel time, which can plausibly benefit both female and male but at a different magnitudes. Standard urban economic theory tells us that this benefit encourage the residents to commute further as well as induces in-migration of workers into the nearby area of MRT stations, which will result in higher work participation rate and housing prices in those areas.[5] The resulting residential relocation itself is hard to analyze due to the data limitation, and 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 complicates the response of male and female labor supply decisions. For example, if a family (couple) faces a reduction of commuting cost by the Metro and a high paid job gets accessible to the husband, one of possible responses is wife's withdrawal from market economy activity (increase home production), substitutively increasing male's labor supply (i.e., intensification of division of labor). Please note that, when we measure outcomes for females relative to males, these confounding mechanisms, other than improved safety, might be also influencing the overall impact.

For the treatment variable, we define the (log) distance to the nearest Phase I or II Delhi Metro station. The reason of this choice of continuous treatment variable follows Gibbons et al. (2017), which suggests to use a continuous treatment intensity (such as distance) as the treatment variable rather than a binary one (connected or not) if the

[5]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 actual signs and magnitude of the impact ambiguous.

transport network in the study area is already dense before intervention. This generally applies to large cities, and Delhi is not an exception where a dense local transport network of railway, bus, and other services had already existed before the introduction of the Delhi Metro.

As described in the previous section, our data is neither experimental nor quasi-experimental. Our unit of observation is aggregated at the level of zones (town or ward), which divide the NCT (National Capital Territory) of Delhi into around 340 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 (1991, 2001) under the common trend assumption. Our DID estimation sets the year 2001 as the baseline year, and treat the 2011 as the end-line. The additional pre-treatment observation, the year 1991, is included as the "lead" (Angrist and Pischke, 2009) period in the manner of Autor (2003), in order to test the common trend. More specifically, we estimate the following equation on the three-period panel,

$$Y_{it} = \alpha_i + \lambda_t + \beta D_{it} + \beta_{-1} D_{it}^{pre} + \delta X_{it} + \epsilon_{it} \tag{2}$$

where, Y_{it} is the outcome variable of zone i at year t; D_{it} is our treatment variable, the log of average distance to nearest Phase I or II metro station. Namely, for the post-treatment observation, $D_{i,2011} = D_i$. For the two pre-treatment period, the value takes $D_{i,2001} = 0$ and $D_{i,1991} = 0$. D_{it}^{pre} is the "lead" of the treatment, which takes $D_{i,1991}^{pre} = D_i$ and zero for other years. This term is included so that we can jointly assess the validity of the common trend assumption in our data. If this term is significant, the treatment assignment predicts the 1991 outcome and indicating the endogenous alignment of metro location in the areas of study. If on the other hand this term is insignificant, the treatment assignment and re-treatment trends are uncorrelated. Given the definition of D_{it} and D_{it}^{pre} above, the equation (2) can be rewritten as

$$Y_{i,2011} = \alpha_i + \lambda_{2011} + \beta D_i + \delta X_{i,2011} + \epsilon_{i,2011} \quad \text{:End-line (post treatment)} \tag{3}$$

$$Y_{i,2001} = \alpha_i + \lambda_{2001} + \delta X_{i,2001} + \epsilon_{i,2001} \quad \text{:Baseline (pre-treatment)} \tag{4}$$

$$Y_{i,1991} = \alpha_i + \lambda_{1991} + \beta_{-1} D_i + \delta X_{i,1991} + \epsilon_{i,1991} \quad \text{:Lead (pre-treatment)} \tag{5}$$

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.[6]; ϵ_{it} is the error term. The coefficient β will capture the treatment effect, and the sign and the magnitude of this coefficient is our central concern. β_{-1} is the coefficient on the "lead" term. We expect that β_{-1} is insignificant under the common trend assumption.

In practice, there are two widely used approaches to estimate equation (2). First approach is the "within" estimation,

$$Y_{i,t} - \bar{Y}_i = \lambda_t - \bar{\lambda} + \beta \left(D_{it} - \bar{D}_i\right) + \beta_{-1} \left(D_{it}^{pre} - \bar{D}_i^{pre}\right) + \delta \left(X_{it} - \bar{X}_i\right) + \epsilon_{it} - \bar{\epsilon}_i \tag{6}$$

Where, $\bar{z}_i$ is time-average of variable z_{it} for individual i. [7]

We estimate each of equation (6) with the set of controls X_{it}[8]. The variables in X_{it}

[6]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 = $\frac{\text{share of scheduled caste in female population}}{\text{share of scheduled caste in male population}}$

[7] Another option is taking the first difference to taking out the fixed effect α_i,

$$\Delta Y_{it} = \Delta\lambda_t + \beta \Delta D_{it} + \beta_{-1} \Delta D_{it}^{pre} + \delta \Delta X_{i,t} + \Delta\epsilon_{it} \tag{7}$$

One potential caveat of (7) is that the error term $\Delta\epsilon_{it}$ is serially correlated by construction. We address this issue by calculating the cluster-robust standard error with clustering at the level of zone. The results with this first-differenced equation are not shown in the paper, while the results are almost same as those with within estimator.

[8]We also conduct estimation without X_{it} and get qualitatively the same results as those with the controls

are household size, child (under six years old) share in population, female-male ratio of literacy rate and scheduled caste share. 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.[9] 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); (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)

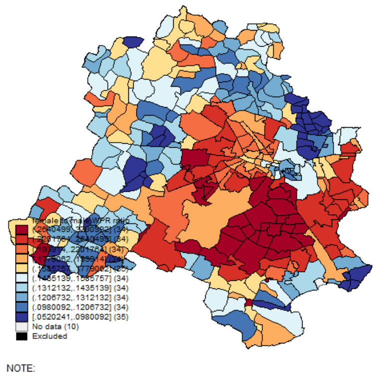
5 Results

Tables 3, 4, 5, and 6, report the results of estimation across different specifications. Table 3 reports the estimation results of equation (6) taking the female WPR as the outcome with location specific time-variant characteristics, X_{it}. For all the five subset analysis, our treatment variable, D_{it}, is significant at 1 percent significance level with negative sings, except for the column (5) with significance at 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 the column (1)

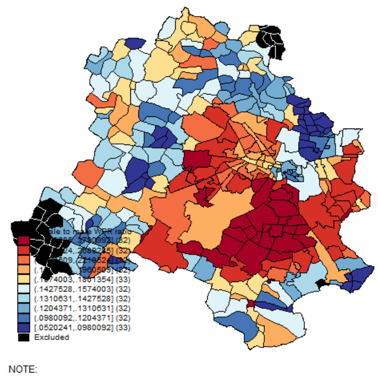
[9] In the separate regression, we check these variables do not seem to be the consequences of the treatment D_{it}, allowing us to included them as controls in the equation.

Figure 5: Subsample Definition and “WPR ratio” in 2011

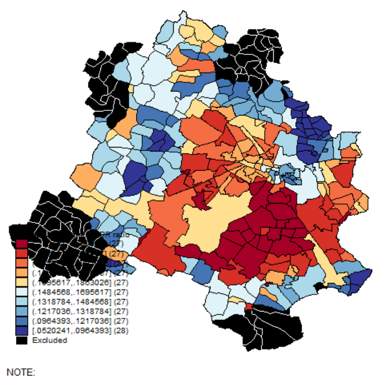
(a) All zones in Delhi (1)



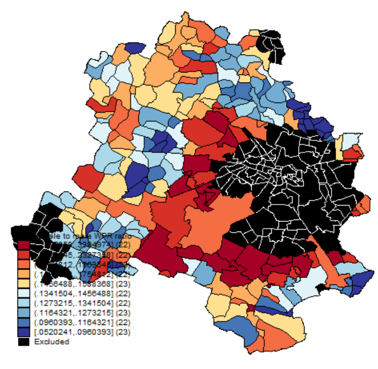
(b) Within 10km reach from commissioned and planned Metro Stations (2)



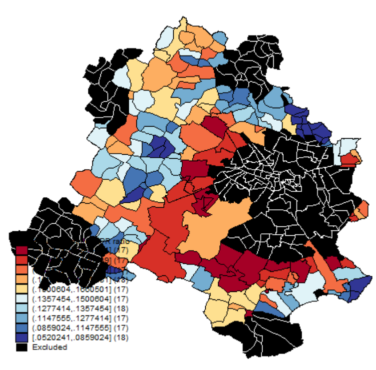
(c) Within 5km reach from commissioned and planned Metro Stations (3)



(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)



of the Table 3, if the distance to the nearest Phase I or II station becomes double, female WPR decreases by 0.558 percentage points. Given that the mean of female WPR in 2011 is 7.91%, this implies that doubling the distance around the mean distance of 5.239km will reduce WPR of female to 7.35%.

The columns (2) and column (3) of Table (3) limit the sample zones within 10 km and 5 km access to any Metro station regardless of whether it has already been commissioned as of 2011 or not (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 the point in time, so they may share the similar pre-treatment unobserved characteristics with zones close to the commissioned stations. By estimating the model of the column (2) and column (3), we compare the outcomes in zones got access to metro stations earlier with those would get it later.

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

The results shown in the Table 3 suggest that a positive effects of the accessibility to the Delhi Metro for female exists. For all the columns, the coefficients on the "lead" term are insignificant, which means these subsets are relatively well defined to ensure the common pre-trend assumption.

Table 4 shows the results for the impact on male WPR. Contrary to the case for females, all the coefficients on distance to commissioned Metro station are positive and significant at 1 percent or 5 percent significance level. The parallel pre-treatment trend assumption is overall satisfied except for the column (3) whose coefficient on the "lead" term is negative and 10 % statistically significant. Furthermore, the magnitude of the effect does not vary across subsamples, ranging from 0.00801 to 0.00975, compared to the case for female shown in Table 3. From the results in Table 3 and Table 4, it turns out that the proximity to the Delhi Metro station affects positively for female WPR while its impact is negative for that of male. Given that the mean of WPR of male in 2011 is 45.3%, this implies that doubling the distance around the mean distance of 5.239km will increase WPR of male to 46.2%.

Table 5 reports the results when the outcome variable is the WPR ratio between female and male. Consistent with the results in Table 3 and Table 4, the coefficients on the distance to commissioned station are negative and significant at 1 percence level. The results implies that the gap of WPR between female and male 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 for this subset as the coefficient of "lead" is insignificant.

Finally, Table 6 reports the results when the total WPR is used as the outcome. Total WPR is the sum of female and male main worker in non-agricultural sector divided by total adult population. For the first three columns show significantly positive coefficients on the distance to nearest commissioned Metro station, meaning that the proximity to Metro station affects negatively to total work participation. However, as shown in the column (4) and column (5), the effect becomes no longer significant suburban subsamples. In the area outside of the CBD premises, proximity to the Metro does not change the overall work participation.

From the results above, we can summarise about the potential impact of Delhi Metro on the work participation as follows. Firstly, females and males have reacted oppositely. Female's WPR in 2011 is higher in zones close to the Delhi Metro station, while it is in the distant zones from the Metro station where male's WPR is higher. Therefore, in areas closer to Metro, it seems that female's economic participation expanded more intensively than that of male's. Secondly, especially for female, the magnitude is larger for the suburb subsamples. This means that the difference caused by the access Delhi Metro might be more pronounced in the suburban area than the CBD premises. Thirdly, partially reflecting that female is positively affected by the proximity and male is negatively affected, the total WPR is negatively affected, because the impact on male surpass that on female, except for the suburban subsamples.

In addition, our results could be suggestive to an emerging literature on labour-leisure choice of married couple in urban context, which are studied using the case of developed countries (e.g. Abe (2011); Black et al. (2014); Johnson (2014); Kawabata and Abe (2018)). The studies have revealed that the labour-leisure choice of married women is substantially different from those of single women and males, and it is closely related with commuting time to CBD(Central Business District)'s. One of the important implication of this literature is the potential reservation wage effect of improved urban transportation system. If the commuting cost for the breadwinner (husband) reduces, it enlarges his effective labour market and earning opportunities. Under certain condition, this may induce wives to consume more leisure (concentrate on household production), despite her effective labour market also expands. For the case of the Delhi Metro Phase I and Phase II, the system has a hub-spoke design from the CBD ("Connaught Place"), making the suburbs on the spokes more accessible to the CBD. Given that, there is a possibility that the Metro induces male-breadwinner of families living in suburbs to find higher paid job in the CBD with making his wife more concentrates on household production by reducing economic activity (work). If this happens, the WPR ratio in zones close to the Metro stations should be lower, especially in the suburbs. Our first round assessment from the current analysis suggests that this reservation wage hypothesis seems not be the case for Delhi because the opposite is observed in our sub-sample analysis.

Table 3: Impact of Proximity to the Delhi Metro on Work Participation Rate of Females (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)	-0.00558***	-0.00688***	-0.00418***	-0.00906***	-0.00453**
	(0.00146)	(0.00166)	(0.00152)	(0.00230)	(0.00207)
Dist. to Metro(lead, 1991)	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

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 Discussion

In this paper, we examine and quantify the impact of Delhi Metro's first two project phases on the work participation rate of females, relative to males. Given the rapid urbanization and the refreshed development goals (SDGs), the urban transportation is not only expected to play a key role in regional aggregate economic growth but also in inclusion of those in vulnerable situations, women, children, persons with disabilities and older persons into the society, hence to improve the social welfare by enhancing their capacity to access to various

Table 4: Impact of Proximity to the Delhi Metro on Work Participation Rate of 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)	0.00862***	0.00993***	0.00975***	0.00801**	0.00829**
	(0.00217)	(0.00225)	(0.00237)	(0.00327)	(0.00329)
Dist. to Metro(lead, 1991)	-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

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

socio-economic opportunities.[10]

Benefiting from a three-period panel from the India's census that provides various demographic information of more than 300 geographical zones within Delhi, we analyse the impact 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

[10]SDGs also emphasizes the inclusiveness in infrastructure investments. For example, Goal 9.1: Develop quality, reliable, sustainable and resilient infrastructure, including regional and trans-border infrastructure, to support economic development and human well-being, with a focus on affordable and equitable access for all.; Goal 11.2: By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons.

Table 5: Impact of Proximity to the Delhi Metro on a ratio of Work Participation Rate of Females over that of 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)	-0.0166***	-0.0210***	-0.0120***	-0.0278***	-0.0160***
	(0.00438)	(0.00476)	(0.00395)	(0.00676)	(0.00542)
Dist. to Metro(lead, 1991)	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

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

female and male. Thanks to the data structure with two pre-treatment period observations, we employ the Difference-in-Differences estimation with "lead", à la Autor (2003), which enable us to verify the common trend assumption during the pre-treatment periods.

The overall results suggest that the proximity to the Metro station have positive effects on female's work participation, while it has worked oppositely for males. This suggests that the Metro probably encouraged females to participate in economic activities more than males, potentially having caused replacement of male by female. However, we still need further investigation to know the mechanisms behind it. More specifically, with

Table 6: Impact of Proximity to the Delhi Metro on 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)	0.00326**	0.00326**	0.00406**	0.00131	0.00289
	(0.00152)	(0.00163)	(0.00160)	(0.00233)	(0.00218)
Dist. to Metro(lead, 1991)	-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

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

current dataset, we cannot tell why the positive effect on female economic participation rather than male is observed. It is unclear whether the improved safety of commuting path encouraged women to take a job outside of their home, since we do not have commuting information. Alternative stories driven by labour demand can generate the same pattern of work participation rate. For instance, the Delhi Metro stations could have stimulated commercial activities around Metro stations, such as retail shops, restaurants, offices, etc. If some female oriented services (either by gender-wage gap or stakeholders' preference/discrimination) flourish in station nearby areas, it would create female employment opportunities more than those for males. In this case, it is not because of the safety of the Metro facility itself, but the type of industries attracted to the premises of the Metro stations, that generate the observed pattern of female and male work participation rate.

Furthermore, in the current analysis, we cannot take into account the female's and male's decision making at the family level, especially the case of married couples. As we discuss above, our results might not support the hypothesis of reservation wage effect of urban transport on wife, which is caused by the breadwinner (husband)'s increased access to better-paid job . However, to assess the existence of this effect for the case of Delhi definitely requires micro-data of married couple.

Finally, current analysis could be prone to the bias arising from the measurement error in the choice of geographical units, as well as the spatial autocorrelation, which we observe in Figure 3. Robustness check with alternative geographical units as well as properly addressing the spatial autocorrelation, by employing spatial econometrics, would be a necessary step to make the inference more trustworthy.

References

Abe, Yukiko, “Family labor supply, commuting time, and residential decisions: The case of the Tokyo Metropolitan Area,” *Journal of Housing Economics*, 2011, *20* (1), 49–63.

African Development Bank Group, “Checklist for Gender Mainstreaming in the Infrastructure Sector,” 2009, (January).

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

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

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

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

Black, Dan A., Natalia Kolesnikova, and Lowell J. Taylor, “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*, 2014, *79*, 59–71.

Borker, Girija, “Safety First: Perceived Risk of Street Harassment and Educational Choices of Women,” 2017, (November).

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

Gibbons, Stephen, Teemu Lyytik, Henry Overman, and Rosa Sanchis-guarner, "New Road Infrastructure: the Effects on Firms," *SERC Discussion Paper*, 2017, (September).

Gimenez-nadal, J Ignacio and Jose Alberto Molina, "Commuting Time and Labour Supply in the Netherlands A Time Use Study," *Journal of Transport Economics and Policy*, 2014, *48* (3), 409–426.

_ , _ , and J Ignacio Gimenez-nadal, "Commuting Time and Household Responsibilities : Evidence Using Propensity Score Matching," 2015, (8794).

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

Goel, Rahul and Geetam Tiwari, "Access-egress and other travel characteristics of metro users in Delhi and its satellite cities," *IATSS Research*, 2016, *39* (2), 164–172.

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

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

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

Johnson, William R, "House prices and female labor force participation," *Journal of Urban Economics*, 2014, *82*, 1–11.

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

Klasen, Stephan and Janneke Pieters, "What explains the stagnation of female labor force participation in Urban India?," *World Bank Economic Review*, 2015, *29* (3), 449–478.

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

Peters, Deike, "Gender and Sustainable Urban Mobility," Technical Report 2013.

Safetipin, "Using Data to Build Safer Cities," Technical Report 2016.

Seki, Mai, "Towards the Inclusive Growth: Literature Review on the Impact Evaluation of Infrastructure (in Japanese)," *JICA-RI Development Cooperation Literature Review Series No.4*, 2016, *March* (4).

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

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

United Nations, *World urbanization prospects* 2014.

Uteng, Tanu Priya, "Gender and Mobility in the Developing World," *World Development Report Background Paper*, 2011, p. 98.

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

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

Appendix

A Other Estimations

Table A.1: Log Total Population (Within Estimator (6), With 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)	-0.0118	0.0861*	0.183***	-0.0632	0.0470
	(0.0479)	(0.0513)	(0.0627)	(0.0767)	(0.0942)
Dist. to Metro(lead, 1991)	0.0486	0.0333	-0.0937	0.312***	0.170*
	(0.0638)	(0.0689)	(0.0853)	(0.0777)	(0.0985)
Household Size	-0.279	-0.0902	-0.257	-0.860	-1.023*
	(0.469)	(0.479)	(0.513)	(0.527)	(0.571)
Children Share	0.581	0.487	0.448	1.310**	1.500**
	(0.529)	(0.549)	(0.728)	(0.563)	(0.668)
female to male literacy ratio	1.226***	1.132***	1.302**	0.701***	0.702**
	(0.374)	(0.371)	(0.522)	(0.219)	(0.336)
female to male SC ratio	0.549	0.195	0.379	0.159	-0.0220
	(0.770)	(0.912)	(1.033)	(0.891)	(1.003)
Constant	11.27***	10.88***	11.46***	13.15***	14.19***
	(1.243)	(1.268)	(1.716)	(1.189)	(1.399)
Year Dummy	YES	YES	YES	YES	YES
Observations	1,006	948	801	654	507
R-squared	0.166	0.190	0.252	0.261	0.327
Number of id	342	322	271	224	173
Adj-R	0.159	0.183	0.245	0.252	0.316

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 A.2: Log Female Population (Within Estimator (6), With 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)	-0.0170	0.0814	0.181***	-0.0695	0.0452
	(0.0481)	(0.0515)	(0.0631)	(0.0766)	(0.0945)
Dist. to Metro(lead, 1991)	0.0492	0.0328	-0.0927	0.312***	0.173*
	(0.0640)	(0.0689)	(0.0851)	(0.0779)	(0.0989)
Household Size	-0.164	0.0351	-0.117	-0.723	-0.867
	(0.480)	(0.489)	(0.524)	(0.542)	(0.587)
Children Share	0.624	0.533	0.456	1.391**	1.533**
	(0.529)	(0.551)	(0.730)	(0.552)	(0.677)
female to male literacy ratio	1.262***	1.165***	1.334**	0.739***	0.742**
	(0.380)	(0.377)	(0.528)	(0.233)	(0.357)
female to male SC ratio	0.415	0.0171	0.222	0.00965	-0.141
	(0.770)	(0.908)	(1.029)	(0.881)	(0.997)
Constant	10.37***	9.973***	10.45***	12.28***	13.19***
	(1.250)	(1.274)	(1.722)	(1.173)	(1.422)
Year Dummy	YES	YES	YES	YES	YES
Observations	1,006	948	801	654	507
R-squared	0.175	0.201	0.263	0.272	0.336
Number of id	342	322	271	224	173
Adj-R	0.169	0.194	0.256	0.263	0.325

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 A.3: Log Male Population (Within Estimator (6), With 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)	-0.00731	0.0902*	0.186***	-0.0578	0.0490
	(0.0477)	(0.0512)	(0.0625)	(0.0767)	(0.0940)
Dist. to Metro(lead, 1991)	0.0481	0.0335	-0.0945	0.312***	0.168*
	(0.0638)	(0.0690)	(0.0856)	(0.0775)	(0.0982)
Household Size	-0.367	-0.186	-0.363	-0.962*	-1.139**
	(0.463)	(0.473)	(0.507)	(0.519)	(0.562)
Children Share	0.554	0.459	0.447	1.260**	1.484**
	(0.529)	(0.550)	(0.728)	(0.571)	(0.663)
female to male literacy ratio	1.202***	1.110***	1.281**	0.675***	0.675**
	(0.370)	(0.368)	(0.518)	(0.211)	(0.323)
female to male SC ratio	0.651	0.330	0.499	0.273	0.0694
	(0.773)	(0.918)	(1.040)	(0.902)	(1.011)
Constant	10.76***	10.38***	11.02***	12.63***	13.75***
	(1.242)	(1.267)	(1.714)	(1.202)	(1.384)
Year Dummy	YES	YES	YES	YES	YES
Observations	1,006	948	801	654	507
R-squared	0.158	0.181	0.244	0.253	0.320
Number of id	342	322	271	224	173
Adj-R	0.152	0.174	0.236	0.243	0.309

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 A.4: Log Total Main Workers, Within Estimator, With 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)	0.00832	0.104*	0.202***	-0.0556	0.0578
	(0.0493)	(0.0531)	(0.0646)	(0.0781)	(0.0958)
Dist. to Metro(lead, 1991)	0.0332	0.0187	-0.0978	0.301***	0.161
	(0.0762)	(0.0823)	(0.102)	(0.0810)	(0.103)
Household Size	-1.380**	-1.146**	-1.267**	-2.068***	-2.163***
	(0.534)	(0.543)	(0.580)	(0.614)	(0.660)
Children Share	0.380	0.268	0.191	1.191**	1.297*
	(0.570)	(0.591)	(0.795)	(0.541)	(0.702)
female to male literacy ratio	1.384**	1.272**	1.598*	0.567**	0.639
	(0.613)	(0.611)	(0.828)	(0.257)	(0.395)
female to male SC ratio	0.616	0.277	0.607	0.251	0.203
	(0.817)	(0.963)	(1.103)	(0.905)	(1.040)
Constant	11.42***	10.94***	11.42***	13.56***	14.37***
	(1.378)	(1.400)	(1.925)	(1.165)	(1.477)
Year Dummy	YES	YES	YES	YES	YES
Observations	1,006	948	801	654	507
R-squared	0.218	0.235	0.296	0.328	0.385
Number of id	342	322	271	224	173
Adj-R	0.211	0.229	0.289	0.319	0.375

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 A.5: Log Female Main Workers, Within Estimator, With 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)	-0.0393	0.0405	0.173**	-0.158*	0.00316
	(0.0526)	(0.0571)	(0.0681)	(0.0831)	(0.0992)
Dist. to Metro(lead, 1991)	-0.0444	-0.0538	-0.146	0.305***	0.224*
	(0.0845)	(0.0902)	(0.110)	(0.102)	(0.125)
Household Size	-2.288***	-2.160**	-2.091**	-2.830**	-2.671**
	(0.861)	(0.916)	(0.934)	(1.120)	(1.165)
Children Share	-0.887	-1.104	-1.167	-0.0564	0.142
	(0.665)	(0.688)	(0.896)	(0.681)	(0.889)
female to male literacy ratio	0.673	0.554	0.931	0.276	0.559
	(0.572)	(0.548)	(0.807)	(0.442)	(0.682)
female to male SC ratio	-0.428	-0.586	-0.0848	-0.383	-0.0943
	(1.025)	(1.151)	(1.321)	(1.122)	(1.316)
Constant	8.146***	7.635***	7.796***	10.18***	10.73***
	(1.638)	(1.691)	(2.148)	(1.702)	(2.031)
Year Dummy	YES	YES	YES	YES	YES
Observations	1,005	947	800	654	507
R-squared	0.341	0.351	0.396	0.438	0.479
Number of id	342	322	271	224	173
Adj-R	0.336	0.346	0.389	0.431	0.470

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 A.6: Log Male Main Workers, Within Estimator, With 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)	0.0217	0.120**	0.212***	-0.0337	0.0708
	(0.0493)	(0.0529)	(0.0644)	(0.0780)	(0.0959)
Dist. to Metro(lead, 1991)	0.0325	0.0174	-0.107	0.304***	0.153
	(0.0749)	(0.0808)	(0.1000)	(0.0812)	(0.103)
Household Size	-1.285**	-1.045**	-1.181**	-1.955***	-2.076***
	(0.522)	(0.529)	(0.567)	(0.585)	(0.631)
Children Share	0.543	0.440	0.347	1.354**	1.438**
	(0.564)	(0.584)	(0.784)	(0.530)	(0.690)
female to male literacy ratio	1.445**	1.332**	1.625**	0.651***	0.693*
	(0.584)	(0.582)	(0.794)	(0.237)	(0.365)
female to male SC ratio	0.744	0.390	0.678	0.383	0.279
	(0.811)	(0.955)	(1.090)	(0.904)	(1.032)
Constant	11.46***	10.99***	11.46***	13.57***	14.38***
	(1.360)	(1.379)	(1.889)	(1.139)	(1.447)
Year Dummy	YES	YES	YES	YES	YES
Observations	1,006	948	801	654	507
R-squared	0.209	0.228	0.287	0.316	0.374
Number of id	342	322	271	224	173
Adj-R	0.202	0.221	0.279	0.308	0.364

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