CDE
(Revised Version) April 2015

DELHI METRO AND AIR POLLUTION

Deepti Goel
Email:dgoel@econdse.org
Department of Economics
Delhi School of Economics

Sonam Gupta
Email: sgupta@impaqint.com
IMPAQ International

(Revised Version, April 2015)

Working Paper No. 229

Centre for Development Economics
Department of Economics, Delhi School of Economics

Delhi Metro and Air Pollution

Deepti Goel and Sonam Gupta*

April 2015

*Deepti Goel is an Assistant Professor at the Department of Economics, Delhi School of Economics, University of Delhi, Delhi 110007. Email: deepti@econdse.org. Sonam Gupta is a Research Associate with IMPAQ International. Email: sgupta@impaqint.com. The authors acknowledge the contribution of the journal's referees in improving the paper and are grateful for their critical feedback. The authors thank Mr. Amit Kumar Jain from the Delhi Metro Rail Corporation, and Dr. D. D. Basu and Mr. D. C. Jakhwal from the Central Pollution Control Board, for their support in providing data used in this paper. They also thank Shantanu Khanna for his excellent research assistance.

Abstract

The Delhi Metro (DM) is an electric based mass rapid transit system serving the National Capital Region of India. It is also the world's first rail project to earn carbon credits under the Clean Development Mechanism of the United Nations Framework Convention on Climate Change for reductions in energy consumption and CO_2 emissions. In this paper we analyze whether the DM led to reductions in localized pollution measured in terms of NO_2, CO, and $PM_{2.5}$, three transportation source pollutants present at dangerously high levels in Delhi. We study the period between 2004 and 2006, and find that one of the larger rail extensions of the DM led to a 34 percent reduction in localized CO at a major traffic intersection in the city. Our study highlights an important social benefit of the metro rail, but it does not advocate the thoughtless building of capital intensive metro rail projects without first undertaking a thorough cost benefit analysis. JEL codes Q5, R4.

1 Introduction

The Delhi Metro (DM) is an electric based mass rapid rail transit system mainly serving the Indian National Capital Territory (NCT) of Delhi. In this paper we examine whether this important mode of public transportation has had any impact on air pollution in Delhi. We identify the immediate localized effect of extending the DM rail network on air pollution measured at two different locations within the city: *ITO*, a major traffic intersection in central Delhi, and *Siri Fort*, a mainly residential neighborhood in south Delhi. Air pollution is measured in terms of three criteria pollutants, namely, nitrogen dioxide **(**NO_2), carbon monoxide (CO), and fine particulate matter **(**$PM_{2.5}$).

An impact study of the DM on air pollution is important for two reasons. First, there is substantial scientific evidence on the adverse effects of air pollution on human health. Block et al. (2012) provide a review of epidemiological research that shows the link between air pollution and damage to the central nervous system which may manifest in the form of decreased cognitive function, low test scores in children, and increased risk of autism and of neurodegenerative diseases such as Parkinson's and Alzheimer's. They also cite studies which show that air pollution causes cardiovascular disease (Brook et al., 2010), and worsens asthma (Auerbach and Hernandez, 2012). Turning to recent research in economics, Ghosh and Mukherji (2014) examine the effect of ambient air quality on children's respiratory health in urban India and find that a rise in particulate matter significantly increased the risk of respiratory ailments. Currie and Walker (2011) find that exposure to vehicular emissions around toll plazas in northeastern United States increased the likelihood of pre-mature births, and also resulted in low birth weight. Some other studies that document the adverse health consequences of air pollution include Moretti and Neidell, 2011; Lleras-Muney, 2010; and Currie, Neidell and Schmieder, 2009.

The second compelling reason for this study is the extent of air pollution in Delhi. According to the World Health Organization's (WHO) database, Ambient Air Pollution 2014, Delhi is the most polluted city in the world in terms of $PM_{2.5}$ levels. In 2013, the annual mean concentration of $PM_{2.5}$ in Delhi was almost twenty times the guideline value prescribed by the WHO.[1] The Central

[1]The WHO guideline for annual mean concentration of $PM_{2.5}$ is 10 $\mu g/m^3$, and in 2013, the annual mean level of

Pollution Control Board (CPCB), the national authority responsible for monitoring and managing air quality in India, finds that pollution in Delhi is positively associated with lung function deficits and with respiratory ailments (CPCB, 2008a and CPCB, 2008b). Guttikunda and Goel (2013) estimate that particulate matter present in Delhi in 2010 led to premature deaths ranging between 7,350 to 16,200 per year, and to 6 million asthma attacks per year. As Delhi continues to grow, population and vehicle densities are bound to increase further, making it all the more important to examine whether the expansion of the DM has had an impact on the city's air quality.

Figures 1A through 1E present the pollution picture at *ITO* during our study period, 2004 to 2006. Each figure shows the 8 or 24 hour average for a specific pollutant along with the corresponding upper limit prescribed by the CPCB. There are some noticeable gaps in each series due to missing observations. In spite of this we see a clear seasonal pattern for nitrogen dioxide (NO_2), carbon monoxide (CO), and particulate matter ($PM_{2.5}$), with their levels being higher in winter (November through January) than in summer (April through June). Further, in case of NO_2, CO, and $PM_{2.5}$, there are a large number of occurrences when their levels exceeded prescribed limits, while there are fewer violations for sulphur dioxide (SO_2) and ozone (O_3). During this period, NO_2, CO, and $PM_{2.5}$, exceeded limits 85, 48 and 78 percent of the time, respectively, while the corresponding figures for SO_2 and O_3 are much lower at 3 and 0.1 percent, respectively.[2] Given that SO_2 and O_3 are within permissible limits most of the time, our analysis focuses on NO_2, CO, and $PM_{2.5}$.

Another reason for restricting focus to only these pollutants is that while NO_2, CO, and $PM_{2.5}$ are mainly generated from transportation sources, SO_2 and O_3 are not. In one of the first pollution inventory studies for Delhi, Gurjar et al. (2004) infer that during their study period (1990-2000), transport sector contributed about 82 percent of nitrogen oxides (NO_x),[3] and 86 percent of CO. In another study for Delhi conducted in 2007, NEERI (2010) reports that the contribution of vehicles

$PM_{2.5}$ in Delhi was 198 $\mu g/m^3$. Notably, Delhi's $PM_{2.5}$ level far exceeded that of Beijing which was at 56 $\mu g/m^3$ (Ambient Air Pollution Database 2014, WHO).

[2]During the same period at *Siri Fort*, NO_2, CO, SO_2 and O_3, exceeded prescribed limits, 13, 26, 0 and 5 percent of the time, respectively. $PM_{2.5}$ was not recorded at *Siri Fort*.

[3]NO_x refers to both, nitrogen monoxide (NO) and nitrogen dioxide (NO_2).

towards NO_x, CO, and particulate matter ($PM_{2.5}$ and PM_{10}), was 18, 58 and 59 percent, respectively. For Delhi in 2010, Guttikunda and Calorie (2012) estimate that 67, 28 and 35 percent of NO_x, CO, and $PM_{2.5}$, respectively, can be attributed to vehicles.[4] While there is variation across studies in the exact share of transportation sources in generating these pollutants, all of them report substantial shares. On the other hand, NEERI, and Guttikunda and Calorie, report that vehicular emissions were responsible for 0.3 percent and 3 percent, respectively, of SO_2.[5] None of these studies look at O_3. However, it is known that O_3 is not directly emitted by motor vehicles, but is created through a complicated non-linear process wherein oxides of nitrogen and volatile organic compounds react together in the presence of sunlight (Sillman, 1999). Thus, of the five pollutants for which we have data, motor vehicles constitute a major and direct source of only three of these, namely, $NO_2,$ CO, and $PM_{2.5}$. To the extent that one of the main channels through which the DM is likely to affect air pollution is through its impact on overall levels of vehicular emissions, we focus our attention on these three pollutants. Moreover, Delhi is a heavily motorized city, and vehicular emissions in particular, is a matter of serious concern.

Theoretical research from transport economics (Vickery, 1969 and Mohring, 1972) postulates the existence of two counteracting effects of introducing a new mode of public transportation on air pollution. On the one hand, introduction of the new mode could increase overall economic activity, which could in turn generate *new* demand for intracity trips. New demand for travel could also be created if the availability of rapid public transport results in a relocation of residents away from the city-center, for example if real estate is cheaper in the suburbs, leading to longer commutes to work. Such demand which did not exist before the new mode was introduced is referred to as the traffic creation effect. If part of the new demand is met by private means of transport, then ceteris paribus this should add to existing levels of vehicular emissions and increase air pollution. On the other hand, with the introduction of a new mode of public transportation, commuters who had earlier relied on private means may now switch to the new mode.[6] This substitution away

[4]The stated contribution of vehicles towards particulate matter in both the preceding studies includes the contribution of road dust.

[5]Gurjar et al. (2004) do not report this figure for SO_2.

[6]According to a report by the Delhi Metro Rail Corporation (DMRC, 2008), the DM has already taken the share

from private to public mode of travel is called the traffic diversion effect. Ceteris paribus, the traffic diversion effect should reduce the overall level of vehicular emissions, and consequently reduce air pollution. In reality both effects are likely to operate. We hypothesize that the traffic diversion effect is likely to dominate the traffic creation effect in the short run. This is because the processes involved in creating new demand for travel are likely to unfold slowly and over a longer period of time, while the traffic diversion effect can occur almost immediately after the new mode is introduced. Nonetheless, it is important to verify this empirically.

To be able to attribute changes in a pollutant measure to the DM, we use the Regression Discontinuity (RD) approach. Our analysis reveals that soon after some of the larger extensions of the DM there were significant reductions in at least some of the transportation source pollutants. Specifically, when we consider our entire study period, 2004-2006, we find that the first extension of the Yellow line, characterized by the largest surge in metro ridership, resulted in a 34 percent reduction in CO at *ITO*. There was also a decline in NO_2 at *ITO* due to the introduction of the Blue line. We are unable to say anything conclusive about $PM_{2.5}$ due to poor quality data on this pollutant.

The rest of the paper is organized as follows. In section 2 we briefly describe the institution of the DM. The empirical strategy is explained in section 3 and the data sources are listed in section 4. Section 5 presents our empirical results. Section 6 ends with policy recommendations.

2 Genesis and Expansion of the Delhi Metro

The Delhi Metro Rail Corporation Limited (DMRC) was set up in 1995 by the governments of Delhi and India to take over the construction and subsequent operation of the DM. Construction work for the metro began in 1998. The first commercial run took place on December 25, 2002, between *Shahdara* and *Tis Hazari* in north Delhi, marking the beginning of operations.

The various stages of expansion of the metro rail network were planned keeping in mind the

of 40,000 vehicles.

expected demand for transportation from different localities. The rail lines were first laid in areas with a high population density, and where it was felt that the metro would benefit the largest number of people. Subsequent extensions were similarly motivated. Table 1 details the phase wise expansion of the DM rail network from its inception in 2002 to 2006. Six extensions were made during our period of study between 2004 and 2006.[7] Figure 2 presents a recent map of the DM rail network. It shows the six extensions made during our study period.[8] Additionally, it also marks the air pollution monitoring stations at *ITO* and *Siri Fort* and the weather station at *Safdarjung*.

3 Empirical Strategy

We use Regression Discontinuity (RD) to estimate the causal impact of the DM on pollution.[9] The basic idea behind this method is explained here. To get at the causal effect we would have ideally liked to compare the levels of pollutants after the metro was extended with their levels, in the same place and at the same time, but in the absence of the metro. However, it is impossible to observe both these scenarios. Therefore, we build the scenario without the metro using observed pollution just before the metro extension. Any sudden change in the levels of pollutants just before and just after the metro extension is attributed to the surge in metro ridership observed at the time of the extension, and is interpreted to be the causal effect of the metro extension. It is important to note that this interpretation is correct only if it were true that in the absence of the metro extension, and after accounting for discontinuous changes due to other known factors such as changing weather conditions, there would have been a smooth transition in the levels of pollutants over time. Later in this section, we talk about the validity of this identifying assumption.

[7] One reason for not studying the period before 2004, is that, we do not have pollution data for it.

[8] Relative to our study period this is a more recent map of the metro network. The note attached to Figure 2 specifies the expansion status of the DM at the end of 2006.

[9] Lee and Lemieux (2010) provide an excellent exposition of this method.

3.1 Estimation Equation

We measure pollution using data from monitoring stations at two different locations within the city, *ITO* and *Siri Fort*.[10] Table 2 presents pollution statistics at each location, along with weather conditions at *Safdarjung*, Delhi. *ITO* has much higher pollution compared to *Siri Fort*: Average hourly NO_2 and CO at *ITO* are 3.2 and 1.5 times their respective levels at *Siri Fort*. This is not surprising given that *ITO* is a major traffic intersection, while *Siri Fort* is a mainly residential area. Ideally, we would have liked to know weather conditions specific to each location. However, we only have hourly weather data for *Safdarjung*, which is fortunately located between *ITO* and *Siri Fort*. We use this as the best available proxy for weather conditions at each location. As the dynamics of pollution is likely to be different across the two locations, and also because they are at different distances from the various line expansions, we estimate impacts at each location separately.

At each location we estimate the impact of a particular metro extension using a time series of hourly pollutant data lying within a symmetric window around that extension's opening date. We also ensure that there are no other extensions within this window. Thus, a window is characterized by a location l (*ITO* or *Siri Fort*), and an extension m. The RD approach is implemented by estimating the following OLS regression within each window:

$$y_t^{l,m} = \theta_0^{l,m} + \theta_1^{l,m} DM_t^m + \boldsymbol{\theta}_2^{l,m}\mathbf{x}_t + \boldsymbol{\theta}_3^{l,m}\mathbf{P}(t) + u_t^{l,m} \tag{1}$$

$y_t^{l,m}$ is pollutant level (in logs) in hour t, at location l, when studying the effect of extension m. DM_t^m is the discontinuity dummy for extension m: Within each window it takes the value 1 for all time periods after the extension date, and 0 for periods before it.[11] $\mathbf{x}_t$, is the vector of covariates

[10] In talking to experts at the CPCB we were told that a monitoring station measures the quality of ambient air passing by it, and it is not possible to demarcate a precise catchment area for which the quality measure would apply. Given that the monitoring stations at *ITO* and *Siri Fort* are approximately 9 kms apart, we believe that they each measure air quality in two distinct geographies within the city. Some evidence for this is provided in table 2 which shows that average pollutant levels are very different across the two locations.

[11] We exclude the 24 hour data pertaining to the day of the extension because we do not know the exact hour when the new line became operational.

and includes controls for weather;[12] for hour of the day, day of the week, and interactions between these two; and for public holidays and festivals such as Diwali.[13] $\mathbf{P}(t)$ is a third-order polynomial in time and captures all smooth variations in pollutant levels. $u_t^{l,m}$ is the error term. The coefficient $\theta_1^{l,m}$ measures the proportionate change in pollutant level at location l as a result of extension m. It is to be interpreted as the immediate localized (at location l) effect on pollution as a result of that particular extension. Since we expect the traffic creation effect to be negligible in the short run, we do not expect $\theta_1^{l,m}$ to be positive. If there is a strong traffic diversion effect, $\theta_1^{l,m}$ will be negative, otherwise it will be insignificant.

Our identification strategy is similar to that used in Chen and Whalley, 2012 (henceforth CW).[14] CW look at the effect of the introduction of the Taipei Metro (TM) on air quality in Taipei City. While they use the discontinuity arising from the *opening* of the metro system, we exploit future discontinuities arising from various *extensions* of the network. Unlike CW, we do not use the first opening of the metro for two reasons. First, we do not have pollution data that dates back to the time when the metro was introduced. Second, even if we had this data, it would be incorrect to use opening ridership discontinuity for Delhi. This is because there was an unprecedented jump in metro ridership when it was first opened and a large part of this jump was due to joy rides which would eventually die out as the novelty of the metro faded away.[15] By using discontinuities in ridership that occur a couple of years after the metro first started, we believe that to a large extent we avoid capturing effects arising from one time rides, and the impact that we measure is closer to the steady state short term effect.

One of the challenges that we faced in estimating equation (1) is the presence of segments of

[12] Controls for weather include current and up to 4-hour lags of temperature, relative humidity, wind speed and rainfall, and quartics of both current and 1-hour lags of these weather variables.

[13] Diwali is a Hindu festival that falls in winter, typically in October or November. It spreads over several days and is celebrated with an ostentatious bursting of firecrackers. It has been documented that air pollution in Delhi shoots up during and immediately following Diwali (CPCB 2012). It is therefore important to control for this source of pollution.

[14] Before Chen and Whalley, Davis (2008) used this method to estimate the effect of driving restrictions on air pollution in Mexico City.

[15] "On the first day itself, about 1.2 million people turned up to experience this modern transport system. As the initial section was designed to handle only 0.2 million commuters, long queues of the eager commuters wishing a ride formed at all the six stations . . . Delhi Metro was forced to issue a public appeal in the newspapers asking commuters to defer joy rides as Metro would be there on a permanent basis."; an excerpt from DMRC, 2008.

missing observations in each pollutant series. The last column of table 2 shows the share of missing observations. The best pollutant series is CO at *ITO* for which 14 percent of observations is missing. $PM_{2.5}$, which is only recorded at *ITO*, has 42 percent missing observations. For the RD strategy to be effective, there cannot be too many missing observations around the extension dates. Therefore, to begin with, we restrict our analysis to only those extensions for which there is a symmetric window of at least nine weeks around the date of extension, wherein missing observations in each included week do not exceed 20 percent of the potential observations.[16] Then we look at other window lengths, and finally, for those pollutants with relatively good data, we analyze the entire series. In order to ensure correct inference in the presence of serial correlation in pollution, in all our specifications we use standard errors clustered at one week.[17]

3.2 Plausibility of Identifying Assumption

Identification of the metro effect breaks down if we have not accounted for an event that has a *discontinuous* effect on air quality.[18] One example is a city wide strike by private bus operators called on the same day as the extension of the metro. If this happens it would be impossible to disentangle the effect of the metro from that of the strike.[19] We have studied the chronology of events in the city and do not find occurrences of such events on any of the extension dates. Here we discuss some of the other likely threats to identification.

Government policies aimed at reducing pollution may have an abrupt impact. One such policy, implemented only in Delhi, was the mass conversion of diesel fueled buses to compressed natural

[16] At the hourly frequency, the number of potential observations in a week is, 168 = 24*7. Each week in our estimation window therefore has at least 134 = 0.8*168 observations.

[17] Although, both Chen and Whalley (2012) and Davis (2008) use standard errors clustered at five weeks, we cluster at one week. This is because our analysis is based on shorter windows of five and nine weeks (due to missing data), while they use two year horizons. Also, for all pollutants in our data, the auto correlation in daily average pollutant level is less than 0.5 beyond seven days. Clustering at one week should therefore be sufficient. Nonetheless, we re-estimated tables 4 and 5 by clustering at two weeks and found similar results.

[18] An event that has a gradual effect on pollution will be captured by the time polynomial and therefore does not impede our analysis.

[19] For them to be problematic, the discontinuos effects do not have to necessarily happen on the extension date. Discontinuous effects arising anywhere within our short windows would be problematic for estimating the correct causal effect of the metro extension.

gas (CNG). However, this happened in 2001, much before our study period began, and is therefore not problematic. In 2005, Delhi moved from Bharat Stage-II to the stricter Bharat Stage-III emission standards. Although this regulatory change was implemented in the middle of our study period, it is unlikely to have led to a sudden change in pollution. This is because the improved norms are only applicable for vehicles manufactured after the new standards are adopted. Given that new vehicle registrations happen uniformly over time, adoption of stricter emission standards should not lead to a sudden drop in vehicular emissions.[20] We do not know of any other regulatory change implemented between 2004 and 2006 that may have had a discontinuous effect on pollution.

Another concern could be that construction activity undertaken to build the new rail lines may have added to localized pollution in the period preceding the metro extension, and this would then over-estimate the DM effect. On speaking to officials from the DMRC we were told that such construction activity is typically completed 15 to 30 days prior to the opening of a new line so as to conduct trial runs to ensure safety of passengers. Therefore, at least for the shorter window lengths, we do not expect this issue to be a problem. Another worry could be that metro officials choose the extension dates in a systematic manner to coincide with either high or low pollution days. We think that this is highly unlikely. Given the public enthusiasm for the metro and the recognition of economies of scale in its operation, the DMRC has always been eager to open a new line once it had met all safety requirements.

Finally, Delhi is characterized by a multitude of pollution sources. According to Guttikunda and Calorie (2012), domestic sources such as burning of bio-fuel for cooking and heating, use of diesel generator sets, waste burning, and construction, together account for 20, 19, and 26 percent of NO_x, CO, and $PM_{2.5}$, respectively. These sources tend to be sporadic, and sometimes mobile, and it is possible that we have not accounted for all of them. In the results section we talk about what definitive conclusions may be drawn in spite of this threat to identification.

[20] Emission standards in India are adopted in a phased manner with stricter norms first being implemented in major cities, including in Delhi, and then extended to the rest of the country after a few years. Given that inter-state freight that plies through Delhi continues to follow the more relaxed emission standards, the impact of Bharat Stage-III within Delhi is dampened.

4 Data Sources

All the data used in this study are from secondary sources. Data on pollutants were obtained from the Central Pollution Control Board (CPCB) which collects it as part of the National Air Quality Monitoring Program (NAMP). We use hourly pollution data recorded at two monitoring stations in the city, namely at *ITO* and at *Siri Fort*.[21] Both are immobile stations that operate on electricity. They provide comparable data as they were bought from the same manufacturer, and followed the same monitoring protocol throughout our study period. Hourly data on weather conditions at *Safdarjung,* Delhi, were obtained from The National Data Center of the India Meteorological Department. Our choice of study period (2004-2006) was dictated by the overlapping period for which we had both pollution and weather data. The Delhi Metro Rail Corporation (DMRC) provided us with data on metro ridership.

5 Results

Before presenting the impact estimates, we investigate whether the data validate a sudden increase in metro ridership at the time of each extension.

5.1 Ridership Discontinuities

For each month, figure 3 shows the percentage change in average daily ridership on the DM over the previous month.[22] The exact magnitudes of change are given in the last column of table 1. Except for the introduction of the Yellow line and the first extension of the Blue line, the figure

[21] Under the NAMP there is one other monitoring station located at the Delhi College of Engineering (*DCE*) in north Delhi. We do not use data from this station because our identifying assumption is unlikely to hold at this location. Compared to *ITO* and *Siri Fort* there are many more erratic sources of pollution at *DCE*. This is because, (a) it is surrounded by Badli, a major industrial township; (b) all along its periphery there are other small scale industrial production units; (c) during our study period the college building itself was undergoing repair and rennovation; and (d) DCE is in a mainly rural part of Delhi where sporadic burning of biomass and wood is widespread.

[22] Actual daily ridership, instead of average daily ridership in a month, would have been ideal in order to check the sudden increase in ridership at each extension date. However, this data was not available for our study period.

shows a significant rise in average daily ridership for the month (or for the following month)[23] of each extension.

The absence of a significant rise in ridership for the introduction of the Yellow line (3 percent increase) may be attributed to the fact that it was the first segment of the north-south corridor, and also a short segment (3 additional stations) that connected the University to the existing Red line at a time when the University was closed for the holiday season. Further, besides the University station, the two other stations on this segment are relatively rich neighborhoods where many people may continue to prefer private over public transportation. For the first extension of the Blue line, the insignificant rise in ridership (5 percent increase) may be attributed to the low population density in south-west Delhi where the extension took place.[24] Given the necessity of observing a large surge in ridership in order to identify the DM effect, we exclude these two extensions from our analysis.

The largest jump in ridership is seen for the first extension of the Yellow line (76 percent increase), which connects areas having a high population density (North-East and Central districts) to the hub of government offices in *Central Secretariat*. A large surge in ridership is also seen for the introduction of the Blue line (56 percent increase) which is the longest extension among all the extensions considered here.

Given this ridership pattern we expect to see larger effects for the first extension of the Yellow line and the introduction of the Blue line. We also expect larger effects at *ITO* than at *Siri Fort* because of its relative proximity to the line expansions, and also because it is a major traffic intersection whereas *Siri Fort* is mainly residential.

[23]For the second extension of the Red line and the first extension of the Blue line, we see the surge in ridership in the month following the one in which the extension took place. This is because these extensions were introduced on the last day of the month, and one would therefore expect average ridership to increase only in the following month.

[24]Of the nine districts of Delhi, the South-West district had the lowest population density of 4,179 persons per square kilometer in 2001. The North-East district had the highest: 29,468 persons per square kilometer (Govt. of NCT of Delhi 2008a).

5.2 Impact of the Delhi Metro

In order to estimate equation (1) using fairly good quality data with fewer missing values, table 3 shows the maximum window length (in weeks) around each extension on applying the at most 20 per cent missing data criteria for each included week, and also subjecting the selection to a minimum window length of five weeks. As an example, if we restrict ourselves to good quality data, we are able to examine the effect of the second extension of the Blue line on NO_2 at *ITO* using a maximum window length of only thirteen weeks. Of the four extensions characterized by a significant increase in ridership, we are unable to examine the effects of the second extension of the Red line because of lots of missing observations around its opening date.

5.2.1 Impact Estimates: Short Windows

Table 4 shows the results from an estimation of equation (1) using a nine week symmetric window of good quality data around each extension date. For each location, it shows the percentage change in the pollutant level that may be attributed to a specific metro extension.[25] Contrary to our expectations, the first extension of the Yellow line did not lead to a statistically significant drop in the level of NO_2 at *ITO*, but as expected it resulted in a huge drop of 69 percent in CO at *ITO*. The introduction of the Blue line resulted in a 31 percent decrease in the level of NO_2 at *ITO*. Its effect on CO at *ITO* could not be analyzed because of missing data. We had expected the second extension of the Blue line to lead to smaller declines and we find that it did not lead to statistically significant reductions in any of the three pollutants. Turning to the effects at *Siri Fort*, we were only able to examine the second extension of the Blue line. Our analysis shows that just as for *ITO*, this extension did not lead to a statistically significant decrease in either NO_2 or CO at *Siri Fort*. It is important to note that even where an effect is not statistically significant, its sign is always negative and in some cases the magnitude is not insignificant.

Table 5 shows the impacts using a shorter window of five weeks. Compared to the nine week

[25] When calculating the percentage change we apply the correction suggested by Kennedy (1981) in the context of interpreting the coefficient on a dummy variable in a semilogarithmic equation.

window, although the magnitude of impact of the first extension of the Yellow line on NO_2 at *ITO* is larger, it is still not statistically significant. The effect on CO at *ITO* for this extension has increased to 78 percent. Also for the introduction of the Blue line, the effect on NO_2 at *ITO* has increased to 55 percent. Restricting to a shorter window enables us to study the effects of this extension on CO and on $PM_{2.5}$: at *ITO* it led to a decrease of 56 and 53 percent, respectively. The results for the second extension of the Blue line at *ITO* and *Siri Fort* are similar to those seen in table 4 and continue to remain statistically insignificant.

As seen in table 3 there are segments of good quality data that span longer than nine weeks. In table 6 we extend the window beyond nine weeks whenever the data permit us to do so. In most cases there is a decrease in magnitude of impact, and none of the effects are significant now. We provide two plausible explanations for the transitory nature of our impact estimates.

One explanation is that some of the sporadic and mobile sources of pollution that characterize Delhi's pollution inventory get captured when we extend the window, and this masks the impacts for longer time periods. Admittedly, this may also happen for shorter windows, and may even explain the very large magnitudes for some of the estimates seen in tables 4 and 5. However, the fact that when we look at shorter time periods we consistently get negative estimates (in table 4 all estimates are negative, and in table 5, all except one, which is close to zero, are negative), makes us believe that some of the larger extensions did reduce specific transportation source pollutants. Alternatively, another explanation could be that the traffic diversion effects are indeed transitory and over longer time horizons the DM has no impact on pollution. Duranton and Turner (2011) provide evidence in support of this argument. They find that in cities in the United States, increase in road building and provision of public transport have no impact on Vehicle-Kilometers-Travelled. They reason that reduced congestion on roads, experienced soon after new roads are built, has a feedback effect which induces existing residents to drive more. If this is true for Delhi, then it is possible that soon after the larger extensions were initiated, the DM diverted private traffic which lowered pollution (as seen in tables 4 and 5 using shorter windows), and also reduced road congestion. These reductions in turn incentivised the remaining drivers to drive more, and may have also added some new drivers, thus wiping out the initial effects on pollution and road congestion (as seen in table 6). This explanation is along the lines of the traffic creation effect discussed earlier. Unfortunately, our data and empirical strategy do not allow us to discern with surety which of these explanations is true. However, our subsequent analysis, using data for the entire study period, suggests that the effects may not be transitory.

5.2.2 Impact Estimates: Using Entire Series

Since our analysis is based on examining particular segments of good quality pollutant data, it is important that observations should not be missing systematically. For each pollutant, Appendix table A1 provides time series of share of missing observations in each month between 2004 and 2006. Eye balling the data does not suggest a pattern to missing observations. We examine this more thoroughly by replacing the outcome variable in equation (1) with an indicator for missing status, and introducing season fixed effects in the regression.[26] Results are presented in appendix tables A2a and A2b. We note that the implementation of extensions does not systematically predict missing status. This is because, for each pollutant, some extensions are positive, some negative, and some insignificant. There is some evidence that observations are less likely to be missing in summer. In order to be certain that our results are not driven by patterns of missing data, we look at the entire series next.

The last column of table 2 shows that CO at *ITO* has the least share of missing observations, 14 percent missing. Therefore, for this series we examine the cumulative effect of several extensions by using all available observations between 2004 and 2006 and estimating the following equation:

$$y_t = \delta_0 + \boldsymbol{\delta}_1 \mathbf{x}_t + \boldsymbol{\delta}_2 \mathbf{P}(t) + \sum_{i=1}^{M} \gamma_i (DMi)_t + \varepsilon_t \tag{2}$$

The variables are similarly defined as in equation (1), and as above, $\mathbf{x}_t$ also includes season fixed effects. The set of discontinuity dummies, $\{DMi\}$, includes all extensions shown in table 3

[26] We were informed by a CPCB official that missing data could be due to one of many reasons: power cuts, instrument failure, software malfunction when transferring data to storage device, and disruption in telephone.

for *ITO*. The cumulative effect of the DM is given by $\{\gamma_i\}$.

Appendix figure A4, visually presents the effects of multiple extensions on CO at *ITO*. The impact magnitudes are presented in table 7 which shows that the first extension of the Yellow line, characterized by the largest increase in ridership, led to a 33.5 percent reduction in CO at *ITO*, while the other two extensions did not lead to statistically significant reductions. Again, it is reassuring that all point estimates are negative even if some are not statistically significant.

We also present similar analysis for NO_2 at *ITO*, which has 18 percent missing observations for the entire series. Figure A5a, which is counterpart of figure A4 for NO_2, seems to suggest that the first extension of the Yellow line and the second extension of the Blue line led to an increase in NO_2 at *ITO*. However, looking at the plot, it seems to us that, unlike CO, the simple time trend does not fully capture the systematic changes for NO_2, and therefore we interact it with the discontinuity dummies. Figure A5b presents the new picture. The effects seen in the previous plot disappear, if anything the introduction of the Blue line seems to have decreased NO_2 at *ITO*.[27] When we estimate equation (2) for NO_2 at *ITO* including the interaction of the discontinuity dummies with the time polynomial, the discontinuity dummy for the second extension of the Blue line drops out, perhaps due to multicollinearity in our dataset. We present the results in Appendix table A3. We note that the coefficient on the introduction of the Blue line is negative and significant.

5.3 Discussion of Results

We start by summarizing our main results and then discuss our estimates in the context of other studies, especially the one by CW.

5.3.1 Summary of Findings

When we look at a short span of nine weeks, we find that the first extension of the Yellow Line, characterized by the largest surge in ridership, led to a 69 percent reduction in CO at *ITO*. When we

[27] This reinforces the result in table 4, in which the introduction of the Blue line was the only extension that showed a significant impact on NO_2 at *ITO*.

extend this span to forty one weeks, the effect size reduces and is no longer statistically significant. However, when we consider our entire study period, 2004-2006, we find that this extension resulted in a 34 percent decline in CO at *ITO*. The fact that we find a decline when we look at the whole series suggests that the effect is not transitory. As pointed out earlier, our identification is not robust to the presence of sporadic and mobile sources of pollution that characterize Delhi. We, therefore, have more faith in our entire period estimate of 34 percent, as disruptive points sources of pollution are likely to be evenly spread within a long window of three years. Moreover, data were not found to be systematically missing over this period.

The introduction of the Blue line, the longest extension considered here, led to a 31 percent reduction in NO_2 at *ITO* when we look at a nine week window, and the effect remains when we consider the entire study period, which once again suggests that this is not a transient effect. Using a five week window, there is some evidence that this particular extension also led to a decline in $PM_{2.5}$ at *ITO*, but we could not carry out the analysis for the entire study period due to a large number of missing observations. Finally, we do not find any significant effects at *Siri Fort* which is mainly a residential area, and relative to *ITO* was further away from the extensions considered here.

In Appendix B, we present two calibration exercises to assess our estimate of a 34 percent reduction in CO vis-à-vis other studies. One exercise suggests that our estimate should be higher, and the other suggests that it should be much lower. We refrain from commenting on the reasons for these differences beyond what we have stated in Appendix B.

5.3.2 Comparison with Chen and Whalley, 2012

CW estimate the impact of introduction of the Taipei Metro (TM), in 1996, on pollution in Taipei City. Using a two-year window of very good quality data (they had 1 missing observation in a two year window), they find that the opening of the TM resulted in a 15 percent decline in CO. Using a three year window and spanning multiple extensions, we find a much larger impact of 34 percent for Delhi.[28] One reason for why our impact is larger could be that CO measurements in our study are from a single monitoring station located at a major traffic intersection, whereas CW use the average CO across ten monitoring stations and they exclude the few stations located at traffic intersections (see footnote 20 on page 15 in CW). If traffic diversion is the main mechanism via which the metro impacts pollution, then one might expect to see bigger effects at traffic intersections.

The pre-metro (1995-96) CO level in Taipei was 1,030 $\mu g/m^3$, while the pre-first-extension-Yellow line CO level for Delhi was 2,212, about twice the level in Taipei.[29] Also, from not having any metro in the city, CW report an average daily ridership of 40,410 in year following the TM introduction. Using data from the DMRC, we note that average daily ridership before and after the first extension of the Yellow line was 119,855 and 385,866, respectively. Thus, a one percentage point decline in CO is associated with an increase in average daily ridership of 2,694 in case of the TM, and an increase of 7,824 for the DM. Given that baseline pollution in Delhi was twice that of Taipei, a one percentage point change in pollution would translate into a much larger absolute reduction for Delhi. Therefore, it is to be expected that Delhi would require a larger change in ridership to support a larger absolute change in pollution. Moreover, a much larger proportion of Delhi's population uses public transport compared to Taipei. For traffic diversion from private vehicles to take effect, Delhi would therefore need a much larger increase in ridership.[30]

[28]Taipei today has a population density of about 15,200 persons per square kilometre, slightly higher than Delhi's 11,050 (Data on population densities were accessed from City Mayors Statistics. Source url: http://www.citymayors.com/statistics/largest-cities-density-125.html. Last accessed on March 31, 2015). According to the inventory study for Taipei City (Chang and Lee, 2006, as cited in CW), 96 percent of CO was due to vehicles. This is comparable to the 86 percent for Delhi reported in Gurjar, 2004, but is much higher than the 58 percent reported in NEERI, 2010.

[29]CW present pollution levels in parts per million (ppm). We converted the figures reported in CW into micro grams per cubic meter ($\mu g/m^3$) using an online convertor provided by Lenntech. Source url: http://www.lenntech.com/calculators/ppm/converter-parts-per-million.htm. Last accessed on March 31, 2015.

[30]The travel mode shares for Taipei City in 2001 (five years after the opening of TM), were 8.8 percent TM, 16.1 percent bus, 34 percent car and 41.1 percent motorbike (Jou et al. as cited in CW). The travel mode shares for Delhi in 2007 (also 5 years since DM opening), were 21 percent walk, 12 percent cycle, 5 percent two-wheelers, 43 percent public transport, 14 percent car and 6 percent three wheelers (Ministry of Urban Development, 2008).

5.3.3 How Viable is the Delhi Metro?

Winston and Maheshri (2007) estimate the contribution of each U.S. urban rail operation to social welfare. They find that with the exception of the BART, the San Francisco Bay area metro system, every system actually reduced welfare. They reason that rail systems are unviable for most U.S. cities because of their high capital costs, declining demand for rail travel, rising labour costs, and inability to raise fares as they have to compete with bus services. The authors report that, on average, rail transit systems in the U.S. cover only about 40 percent of their operating costs. This should raise concerns about the viability of the DM, pollution benefits notwithstanding.

Appendix table A4 shows the annual profits for the Delhi Metro Rail Corporation Ltd. from 2004 to 2010. It is heartening to note that the company recorded positive profits from traffic operations throughout this period, and positive profits from all operations in most years. Delhi has a growing population and high rates of economic growth, and this should alleviate concerns around low ridership demand.[31] As of 2012, the total route length of the DM was 190 kilometers, and its ridership was about 2 million per day (DMRC 2012). In contrast, the San Francisco/Oakland area, has a population density that is slightly more than one fourth that of Delhi's, and in 2014, the BART's route length and average daily ridership was 167 kilometers and 0.4 million, respectively.[32] Moreover, Delhi is a polycentric city and the route design of the DM is ideal to serve daily commuters. It has a radial track layout, with major north-south and east-west corridors connecting different parts of Delhi to the border cities such as Gurgaon and Noida, which are the new centers of employment. Also, ridership is likely to increase further, as new routes get completed. According to Litman (2014) rail transit is more appropriate in areas where development is more compact, and noise and air pollution are serious considerations, while bus is more appropriate where travel is more dispersed. It seems to us that Delhi meets the criteria that favor having a metro. For these

[31]The decadal growth in its population from 2001 to 2011 was 20.96 percent, and for the period 2007-2011, the annual compound growth of its Gross State Domestic Product was 10.1 percent (Govt. of NCT of Delhi 2012).

[32]Data on population density for San Francisco was accessed from City Mayors Statistics. Source url: http://www.citymayors.com/statistics/largest-cities-density-125.html. Route length and daily ridership on the BART was from http://www.bart.gov/sites/default/files/docs/2014BARTFactsheet_Final%20011614.pdf. Both sites were last accessed on March 31, 2015.

reasons we are optimistic about the viability of the DM.[33]

6 Concluding Remarks

We study the impact of the Delhi Metro (DM) on pollution in Delhi between 2004 and 2006. We find that soon after some of the larger extensions of the DM, there were significant reductions in two transportation source pollutants, namely, carbon monoxide and nitrogen dioxide, at a major traffic intersection in central Delhi. Although we find a favorable impact of the DM on Delhi's pollution, the overall impact of the metro system on air pollution crucially depends on how the electricity needed to drive the metro is generated. If it is not cleanly generated then some or all of the benefits from reduced pollution in Delhi may be offset by increased pollution elsewhere.[34]

Much of our analysis was restricted by the poor quality of data on pollution. A longer time series, with fewer missing observations, would have allowed us to draw more definitive conclusions. Given the severity of Delhi's pollution problem, we would urge the environment ministries at the state and the center to invest in better technology and equipment in order to record pollution levels more accurately and completely. As our analysis reveals it is difficult to conduct rigorous impact assessments without good quality data.

Our paper also highlighted the nature of pollution in Delhi by citing several inventory studies. We believe that a multi-pronged strategy needs to be adopted if Delhi wants to shed its distinc-

[33] Murty et al. (2006) carry out a social cost benefit analysis of Phases I and II of the DM, covering a track length of 108 kilometers. Considering the estimates of financial flows during the period 1995-2041, they estimate the financial benefit-cost ratio to be between 1.92 to 2.30. They also estimate the capital costs of Phases I and II to be 64,060 million and 80,260 million, respectively, and the Net Present Social Benefit from both phases to be 419.98 billion INR (in 2004-05 prices). Their calculations account for differences in shadow and market prices of unskilled labour; premiums for importing fuel; benefits accruing from reduced road congestion, accidents, and air pollution; and effects of re-distribution of income among stakeholders. Their estimates are based on several assumptions regarding annual flows of costs and benefits during the entire life time of the DM. Evaluating the accuracy of these assumptions is beyond the scope of our study.

[34] The DM has a regenerative breaking system on its rolling stock, which generates electricity when brakes are applied and then feeds it back into the system. This way, almost 35 percent of the electricity consumed is regenerated by the system (Sreedharan, 2009). Doll and Balaban (2013) provide an excellent analysis of the overall carbon footprint of the DM. For the year 2011, they estimate that the DM saved 232,162 tonnes of CO_2 because of its regenerative braking technology.

tion of being the most polluted city in the world (based on WHO's Pollution Database: Ambient Air Pollution 2014). As potential pollution abatement measures, Delhi's planners should consider the following: increasing accessibility to metro stations by improving feeder systems, promoting cycle-rickshaws to deliver last mile connections, extending the Delhi Bus Rapid Transit System (BRTS) and ensuring strict enforcement of existing bus corridors, using congestion pricing and city-wide parking policies to dissuade use of private vehicles, adopting uniform emissions standards throughout the country, constructing a by-pass road around Delhi to eliminate inter-state freight traffic, designing efficient waste disposal systems to prevent sporadic burning of garbage and foliage, shutting down the remaining coal based power plants in the city, and educating Delhi's residents of the severity of the problem to enable them to take more informed decisions.

Finally, before investing in a capital intensive rail network, it is imperative that a proper cost benefit analysis be undertaken. A specific consideration that needs to be made is its desirability vis-à-vis a bus transit system (BTS). In the context of the United States, Winston and Maheshri (2007), and O'Toole (2010), claim that the less capital intensive BTS is more suitable for most U.S. cities where there isn't enough demand for the metro to be able to recoup its high capital costs. Litman (2014) makes contentions in favour of the metro stating that while it is more capital intensive, it has lower operating costs per passenger mile. Both sides, however, agree that metro systems are best suited in areas with high population density, characterized by high demand for travel. Today, governments at various levels in India are planning to build metro rail systems in several Tier II cities. These cities have lower population densities compared to Tier I cities such as Delhi. Governments, at times, invest in capital intensive projects like the metro to emulate the more advanced economies, or to pander to private firms who build and design these systems. We would caution them against these temptations. As many cities in the United States are realizing, once a metro is built it is hard to abandon it (Winston and Maheshri, 2007).

REFERENCES

Auerbach A., Hernandez M. L. The Effect of Environmental Oxidative Stress on Airway Inflammation. Current Opinion in Allergy and Clinical Immunology 2012; 12(2): 133-139.

Block M. L., Elder A., Auten R. L., Bilbo S. D., et al. The Outdoor Air Pollution and Brain Health Workshop. NeuroToxicology 2012; 33(5): 972-84.

Brook R. D., Rajagopalan S., Pope – III A. C., Brook J. R., et al. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. Circulation 2010; 121: 2331-2378.

Chen Y., Whalley A. Green Infrastructure: The Effects of Urban Rail Transit on Air Quality. American Economic Journal: Economic Policy 2012; 4(1): 58-97.

CPCB (Central Pollution Control Board). 2008a. "Epidemiological Study on Effect of Air Pollution on Human Health (Adults) Delhi." Environmental Health Series 1.

CPCB (Central Pollution Control Board). 2008b. "Study on Ambient Air Quality, Respiratory Symptoms and Lung Function of Children in Delhi." Environmental Health Series 2.

CPCB (Central Pollution Control Board). 2012. "Ambient Air and Noise Pollution Levels - Deepawali 2012." CPCB Press Release, November 13, 2012.

Currie J., Neidell M. J., Schmieder J. F. Air Pollution and Infant Health: Lessons from New Jersey 2009; Journal of Health Economics; 28(3): 688-703.

Currie J., Walker R. Traffic Congestion and Infant Health: Evidence from E-Z Pass. American Economic Journal: Applied Economics 2011; 3(1): 65-90.

Davis L. W. The Effect of Driving Restrictions on Air Quality in Mexico City. Journal of Political Economy 2008; 116(1): 38- 81.

DMRC (Delhi Metro Rail Corporation). 2008. "A Journey to Remember." Public Relations Department, Delhi Metro Rail Corporation Ltd., New Delhi, India.

DMRC (Delhi Metro Rail Corporation). 2012. "Annual Report 2011-12." Delhi Metro Rail Corporation Limited, New Delhi.

Doll N.H.C., Balaban O. A Methodology for Evaluating Environmental Co-Benefits in the Transport Sector: Application to the Delhi Metro. Journal of Cleaner Production 2013; 58(1): 61–73.

Duranton G., Turner M. A. The Fundamental Law of Road Congestion: Evidence from US

Cities. American Economic Review 2011; 101(October): 2616–2652.

Ghosh A., Mukherji A. Air Pollution and Respiratory Ailments among Children in Urban India: Exploring Causality. Economic Development and Cultural Change 2014; 63(1): 191-222.

Goel R., Guttikunda S. K., Mohan D., Tiwari G. Benchmarking Vehicle and Passenger Travel Characteristics in Delhi for On-road Emissions Analysis. Travel Behaviour and Society 2015; 2: 88-101.

Govt. of NCT of Delhi. 2008a. "Economic Survey of Delhi 2008-2009." Planning Department, Government of NCT of Delhi.

Govt. of NCT of Delhi. 2008b. "Statistical Abstract 2008." Directorate of Economics and Statistics, Government of NCT of Delhi.

Govt. of NCT of Delhi. 2012. "Economic Survey of Delhi 2012-2013." Planning Department, Government of NCT of Delhi.

Gurjar B. R., Aardenne J. A., Lelievelda J., Mohan M. Emission Estimates and Trends (1990-2000) for Megacity Delhi and Implications. Atmospheric Environment 2004; 38(33): 5663-5681.

Guttikunda S., Goel R. Health Impacts of Particulate Pollution in a Megacity- Delhi, India. Environmental Development 2013; (6): 8-20.

Guttikunda S., Calori G. Multi-Pollutant Emissions Inventory for the National Capital Region of Delhi. 2012. SIM-air Working Paper No. 38.

Kennedy P. E. Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations. American Economic Review 1980; 71(4), 801.

Lee D., Lemieux T. Regression Discontinuity Designs in Economics. Journal of Economic Literature 2010; 48:281-355.

Litman T. 2014. "Evaluating Rail Transit Criticism." Victoria Transport Policy Institute.

Lleras-Muney A. The Needs of the Army: Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Children's Health. Journal of Human Resources 2010; 45(3): 549-590.

Ministry of Urban Development. 2008. “Study on Traffic and Transportation Policies and

Strategies in Urban Areas in India.” Final Report.

Moretti E., Neidell M. J. Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles. Journal of Human Resources 2011; 46(1): 154-175.

Mohring H. Optimization and Scale Economies in Urban Bus Transportation. American Economic Review 1972; 62(4): 591-604.

Murty M. N., Dhavala K. K., Ghosh M., Singh R. Social Cost-Benefit Analysis of Delhi Metro. 2006. Munich Personal RePEc Paper No. 14465.

NEERI (National Environmental Engineering Research Institute). 2010. "Air Quality Monitoring, Emission Inventory, and Source Apportionment Studies for Delhi." December 2010.

O’Toole R. Defining Success: The Case Against Rail Transit. Policy Analysis 2010; 663:1-40.

Sillman S. The Relation between Ozone, NOx and Hydrocarbons in Urban and Polluted Rural Environments. Atmospheric Environment 1999; 33: 1821- 1845.

Sreedharan E. Climate Change - An Opportunity for Sustainable Development: The DMRC Experience. RITES Journal 2009; 81-88.

TRW (Transport Research Wing). 2013. "Road Transport Year Book (2011-12)." Ministry of Road Transport and Highways, Government of India, New Delhi.

Vickery W. Congestion Theory and Transport Investment. American Economic Review 1969; 59(2): 251-60.

Winston C., Maheshri V. On the Social Desirability of Urban Rail Transit Systems. Journal of Urban Economics 2007; 62: 362-382.

Delhi Metro and Air Pollution

Tables Figures and Appendices A and B

Deepti Goel and Sonam Gupta

April 2015

TABLES AND FIGURES

Table 1 Phase wise extension of the Delhi Metro since inception

S.No.	**Extension**	**Segment (number of stations added)**	**Date of Ext. dd-mm-yyyy**	**Ridership Change (%)[1]**
		Extensions before our study period, 2002-2003		
	Red Line (introduction)	Shahdara-Tis Hazari (5)	25-12-2002	
	Red Line (first extension)	Tis Hazari-Inderlok (4)	04-10-2003	
		Extensions during our study period, 2004-2006		
1	Red Line (second extension)	Inderlok-Rithala (8)	31-03-2004	46
2	Yellow Line (introduction)	Vishwavidyalaya-Kashmere Gate (3)	20-12-2004	3
3	Yellow Line (first extension)	Kashmere Gate-Central Secretariat (6)	03-07-2005	76
4	Blue Line (introduction)	Barakhamba-Dwarka (21)	31-12-2005	56
5	Blue Line (first extension)	Dwarka- Dwarka Sector 9 (6)	01-04-2006	5
6	Blue Line (second extension)	Barakhamba-Indraprastha (3)	11-11-2006	15

Source: Delhi Metro Rail Corporation, DMRC

1. Ridership Change refers to the percentage change in average daily ridership for the month in which the extension was introduced. For example, average daily ridership in March 2004 was 46 percent higher than that in February 2004.

Table 2 Descriptive Statistics for Pollution and Weather, 2004 to 2006

	Mean	Std. Dev.	Min.	Max.	Missing Obs.(%)[1]
ITO (in micrograms per cubic meter)					
NO2	145.5	109.3	0.226	1,140	18
CO	2,389.5	2,328.5	20.6	25,000	14
PM2.5[2]	144.6	131.8	1.69	1,020	42
Siri Fort (in micrograms per cubic meter)					
NO2	45.0	52.5	0.771	805	27
CO	1,632.0	1,818.7	2.98	20,400	52
Safdarjung					
Temperature (deg. C)	25.3	7.9	3.1	43.7	0.01
Relative Humidity (%)	61.3	22.0	5	100	0.0
Wind Speed (kmph)	4.7	5.8	0	62	0.02
Rainfall (mm)	0.07	1.0	0	56	0.0

Source: Authors' calculations using data provided by CPCB and the India Meteorological Department.
1. Missing Obs. refers to percentage of missing observations in the corresponding hourly series.
2. Data for PM2.5 was only recorded at ITO, and only from November 2004 onwards.

Table 3 Window length (in weeks) of good quality data around each extension

Each included week has no more than 20% missing observations
At least five such weeks (symmetric window around extension date)

	NO2	CO	PM2.5
ITO			
Yellow Line (first extension), Jul 03, 2005	19	41	-
Blue Line (introduction), Dec 31, 2005	13	5	5
Blue Line (second extension), Nov 11, 2006	13	13	7
Siri Fort			
Blue Line (second extension), Nov 11, 2006	13	13	-

Source: Authors' calculations using data provided by CPCB.

Table 4 Effect of Delhi Metro on Air Quality: Nine Week Symmetric Window

	NO2	CO	PM2.5[1]
	(Percentage change in level of pollutant)		
ITO			
Yellow Line First Extension (Jul 03, 2005)	-6.6	-69.4***	
Std. Error	(16.5)	(10.5)	
Number of Observations	1457	1497	
Blue Line Introduction (Dec 31, 2005)	-30.6**		
Std. Error	(9.4)		
Number of Observations	1639		
Blue Line Second Extension (Nov 11, 2006)	-10.4	-13.1	-12.4
Std. Error	(9.8)	(8.2)	(16.0)
Number of Observations	1605	1605	1268
Siri Fort			
Blue Line Second Extension (Nov 11, 2006)	-25.9	-3.0	
Std. Error	(24.0)	(9.9)	
Number of Observations	1601	1532	

1. Given the importance of analyzing PM2.5 we report results for a shorter window of 7 weeks for it. Missing observations in each included week do not exceed 20 percent. Each estimate is calculated from a separate regression (equation (1)) where the dependent variable is the natural logarithm of hourly pollutant. Controls are third order polynomial in time; hour of the day, weekday and interaction between the two; current and up to 4-hour lags of temperature, humidity, wind speed and rainfall, and quartics of current and 1-hour lags of the same weather variables; and dummy variables for public holidays and festivities such as Diwali. Std. Errors are clustered at one week. * indicates significantly different from zero at 10 percent level, ** at 5 percent level and *** at 1 percent level.

Table 5 Effect of Delhi Metro on Air Quality: Five Week Symmetric Window

	NO2	CO	PM2.5
	(Percentage change in level of pollutant)		
ITO			
Yellow Line First Extension (Jul 03, 2005)	-21.4	-77.8***	
Std. Error	(21.3)	(6.7)	
Number of Observations	748	772	
Blue Line Introduction (Dec 31, 2005)	-55.2**	-56.0*	-52.8***
Std. Error	(7.7)	(22.8)	(5.3)
Number of Observations	848	848	840
Blue Line Second Extension (Nov 11, 2006)	-6.4	-10.1	-15.8
Std. Error	(5.8)	(8.7)	(15.8)
Number of Observations	935	935	932
Siri Fort			
Blue Line Second Extension (Nov 11, 2006)	-14.2	0.3	
Std. Error	(10.5)	(6.2)	
Number of Observations	934	871	

Missing observations in each included week do not exceed 20 percent. Each estimate is calculated from a separate regression (equation (1)) where the dependent variable is the natural logarithm of hourly pollutant. Controls are the same as used in Table 4. Std. Errors are clustered at one week. * indicates significantly different from zero at 10 percent level, ** at 5 percent level and *** at 1 percent level.

Table 6 Effect of Delhi Metro on Air Quality: Longer than Nine Week Symmetric Window

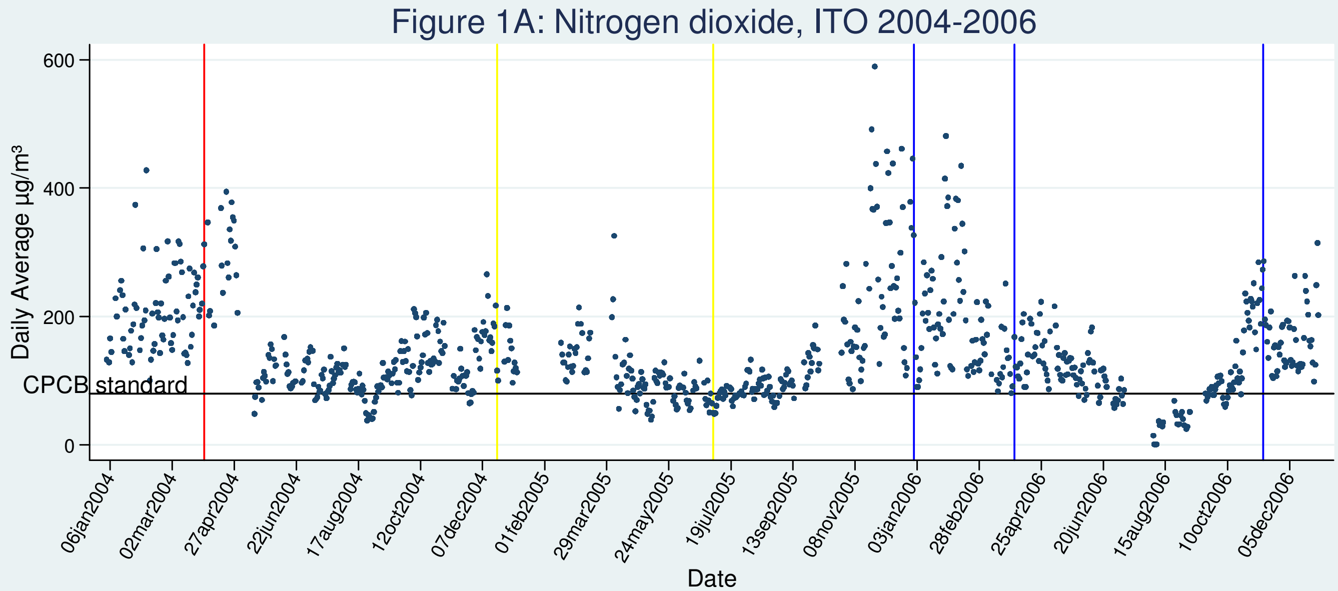
	NO2	CO
	(Percentage change in level of pollutant)	
ITO		
Yellow Line First Extension (Jul 03, 2005)	-5.5	-18.8
Std. Error	(11.5)	(18.7)
No. of Obs. (Window Length in Weeks)	3259 (19)	7118 (41)
Blue Line Introduction (Dec 31, 2005)	-3.1	
Std. Error	(20.5)	
No. of Obs. (Window Length in Weeks)	2281 (13)	
Blue Line Second Extension (Nov 11, 2006)	-2.3	-12.2
Std. Error	(7.6)	(10.5)
No. of Obs. (Window Length in Weeks)	2397 (13)	2397 (13)
Siri Fort		
Blue Line Second Extension (Nov 11, 2006)	-28.4	9.2
Std. Error	(29.4)	(16.2)
No. of Obs. (Window Length in Weeks)	2392 (13)	2315 (13)

Missing observations in each included week do not exceed 20 percent. Each estimate is calculated from a separate regression (equation (1)) where the dependent variable is the natural logarithm of hourly pollutant. Controls are the same as used in Table 4. Std. Errors are clustered at one week. * indicates significantly different from zero at 10 percent level, ** at 5 percent level and *** at 1 percent level.

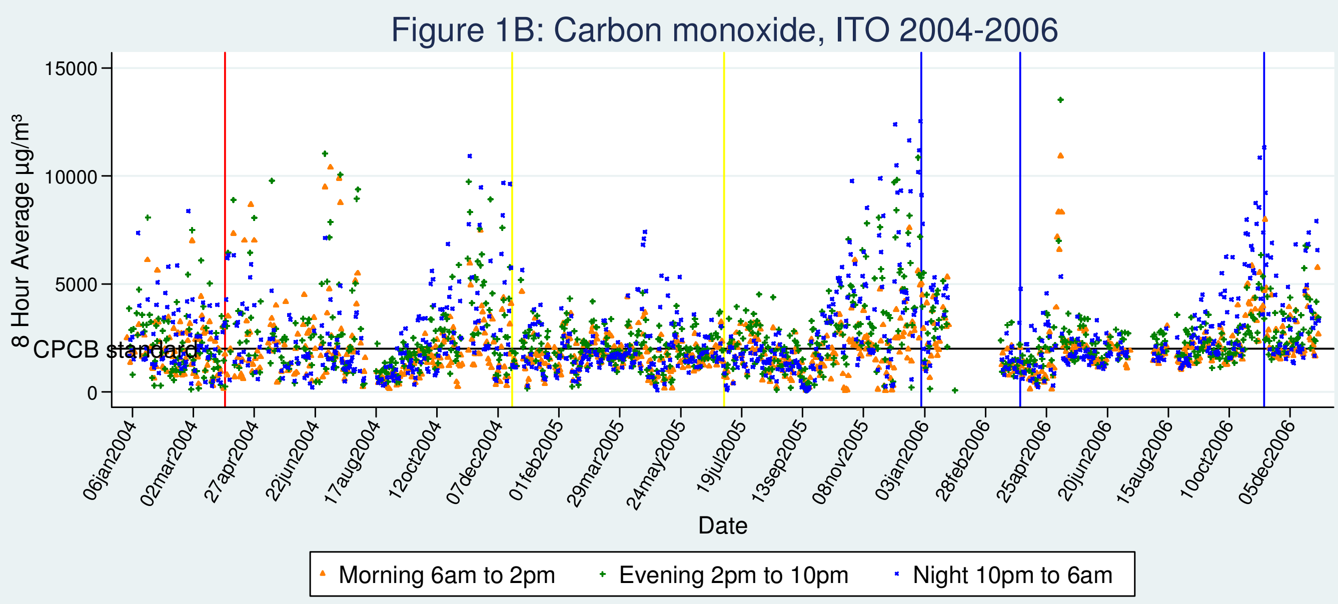
Table 7 Effect of Delhi Metro on CO at ITO: Entire Series, 2004-2006

Metro Extension	**CO at ITO**
Yellow Line First Extension (Jul 03, 2005)	-33.5**
Std. Error	(10.4)
Blue Line Introduction (Dec 31, 2005)	-16.3
Std. Error	(16.6)
Blue Line Second Extension (Nov 11, 2006)	-2.9
Std. Error	(14.6)
Number of Observations	22,158

The dependent variable is the natural logarithm of hourly CO at ITO. Control variables are the same as used in table 4, with the addition of season fixed effects. Std. Errors are clustered at one week. * indicates significantly different from zero at 10 percent level, ** at 5 percent level, and *** at 1 percent level.

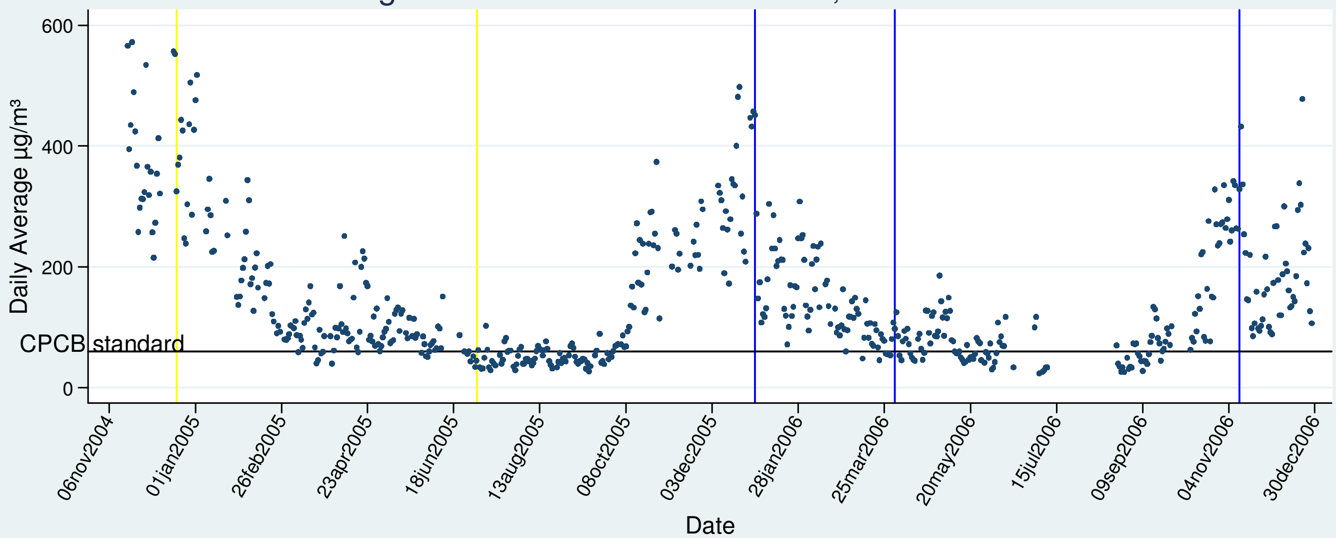


Note: Vertical lines indicate the dates of Delhi Metro expansion
Source: Authors' analysis using pollutant data from the Central Pollution Control Board



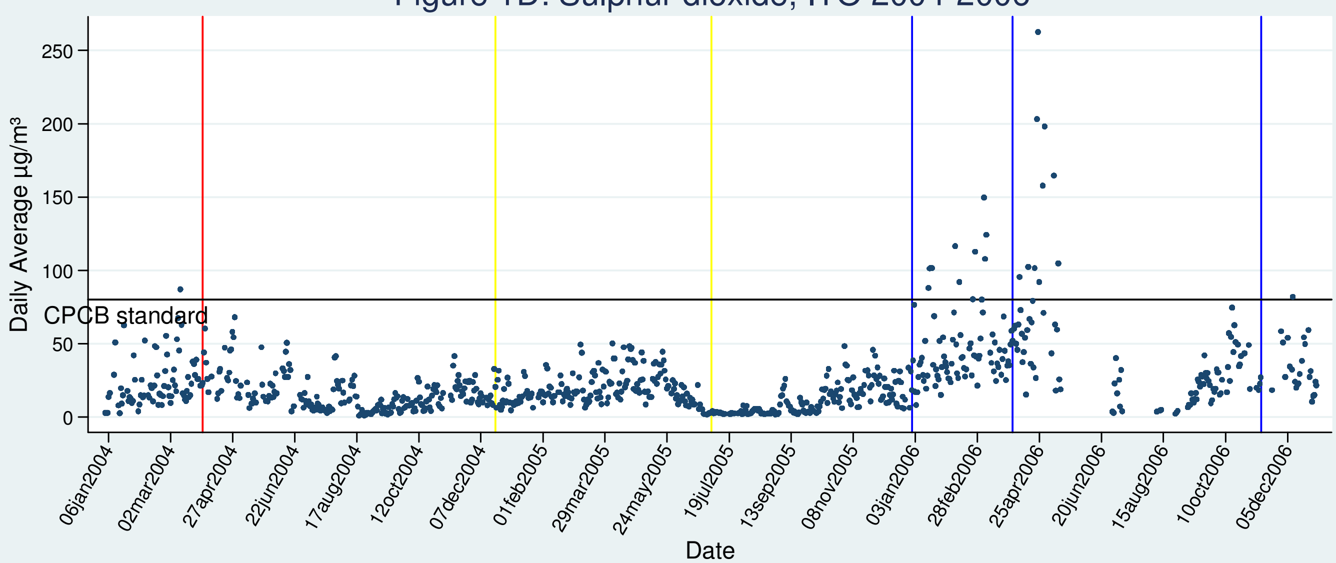
Note: Vertical lines indicate the dates of Delhi Metro expansion
Source: Authors' analysis using pollutant data from the Central Pollution Control Board

Figure 1C: Particulate Matter 2.5, ITO 2005-2006



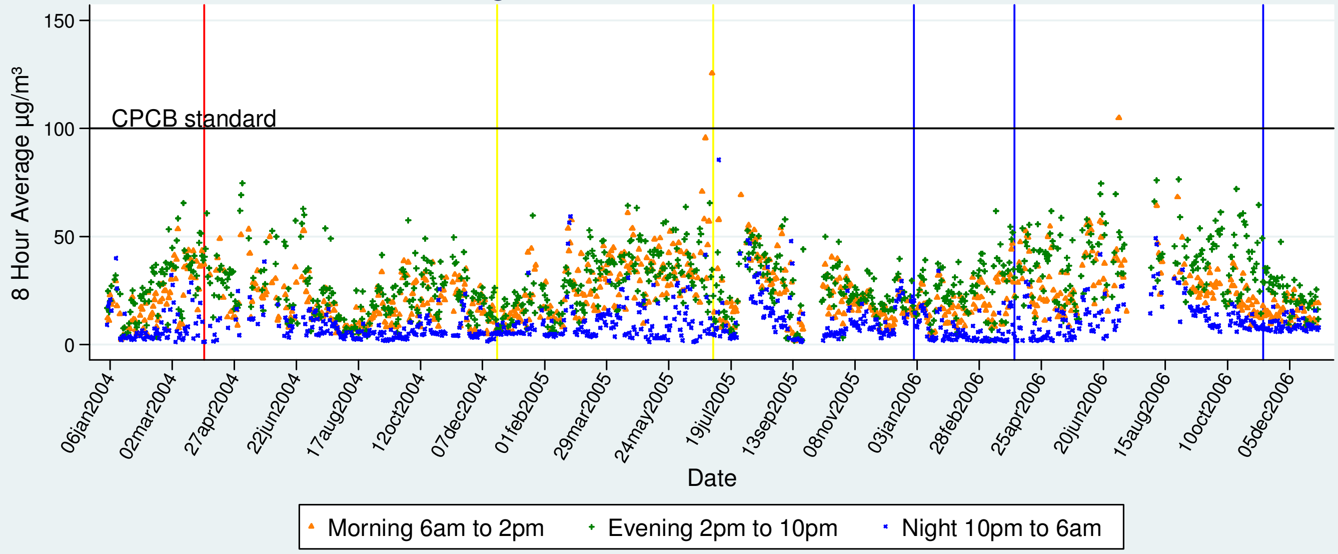
Note: Vertical lines indicate the dates of Delhi Metro expansion
Source: Authors' analysis using pollutant data from the Central Pollution Control Board.
Data for PM 2.5 only available from November 2004 onwards

Figure 1D: Sulphur dioxide, ITO 2004-2006



Note: Vertical lines indicate the dates of Delhi Metro expansion
Source: Authors' analysis using pollutant data from the Central Pollution Control Board

Figure 1E: Ozone, ITO 2004-2006

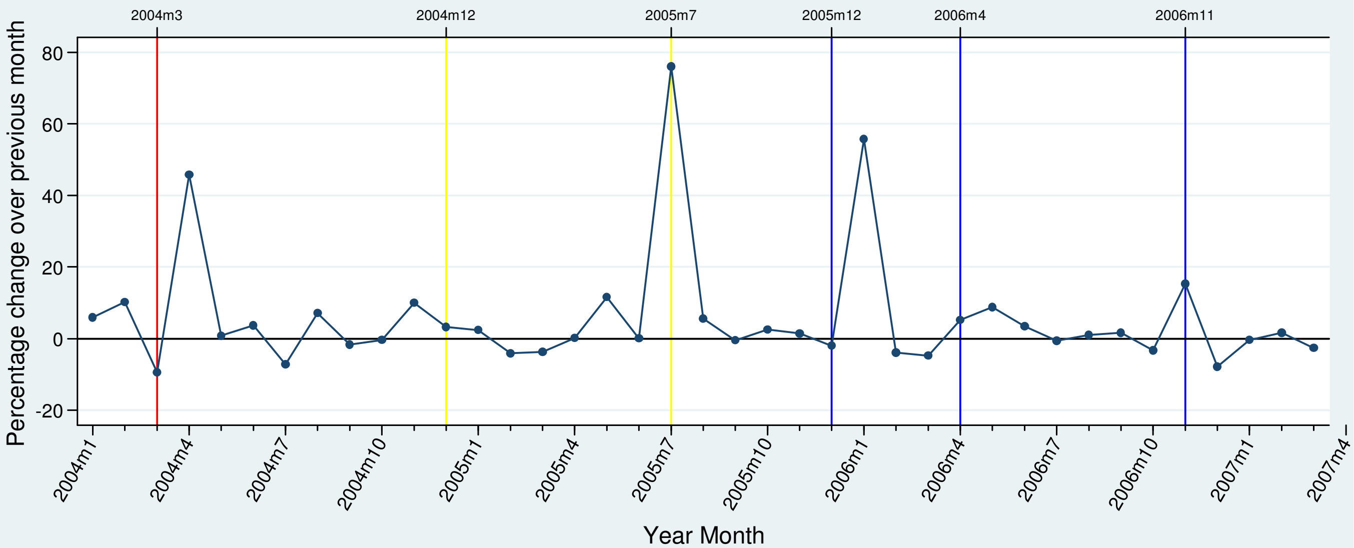


Note: Vertical lines indicate the dates of Delhi Metro expansion
Source: Authors' analysis using pollutant data from the Central Pollution Control Board

Figure 2: Delhi Metro Rail Network and Monitoring Stations

Relative to our study period (2004-2006), this is a more recent map of the metro network. By the end of 2006, the Red line only extended from *Rithala* to *Shahadara*, the Yellow line from *Vishwa* to *Central Sec*, and the Blue line from *Indraprasth* to *Dwarka Sec 9*. The two air pollution monitoring stations are at *ITO* and *Siri Fort*, and are approximately 9 kilometers apart. Relative to *ITO*, *Siri Fort* was not close to any of the extensions that existed during our study period. The weather station is at *Safdarjung*. The star symbols represent the three coal based thermal power plants located within the National Capital Territory of Delhi at the time (one of them was closed later on).

Figure 3: Change in Average Daily Metro Ridership over Previous Month, 2004-2006



Note: Vertical lines indicate the dates of Delhi Metro expansion
Source: Ridership data from the Delhi Metro Rail Corporation

APPENDIX A

Table A1 Month Wise Share (in percent) of Missing Observations, 2004-2006

Year	Month	Extension Made	NO2	CO	PM2.5	NO2	CO
				ITO		Siri Fort	
2004	Jan		20	18	100	54	54
2004	Feb		**8**	**4**	100	**0**	**1**
2004	**Mar**	**Red Line Second Ext.**	**1**	**1**	100	24	24
2004	Apr		30	20	100	42	31
2004	May		47	31	100	28	26
2004	Jun		26	27	100	13	12
2004	Jul		**2**	**1**	100	20	19
2004	Aug		**9**	40	100	13	**5**
2004	Sep		**0**	**0**	100	25	**3**
2004	Oct		**1**	**3**	100	25	**0**
2004	Nov		18	17	56	**3**	**0**
2004	**Dec**	**Yellow Line Intro.**	14	13	23	**5**	**3**
2005	Jan		77	11	44	17	38
2005	Feb		52	**0**	12	**0**	60
2005	Mar		55	**1**	**8**	36	32
2005	Apr		**6**	**0**	**0**	34	50
2005	May		**5**	**3**	**8**	46	100
2005	Jun		11	**6**	33	20	100
2005	**Jul**	**Yellow Line First Ext.**	**4**	**4**	**4**	14	99
2005	Aug		**4**	**3**	**8**	17	100
2005	Sep		26	**1**	**8**	27	99
2005	Oct		56	**3**	**5**	40	100
2005	Nov		**2**	**2**	34	51	100
2005	Dec	**Blue Line Intro.**	**6**	**6**	24	44	100
2006	Jan		**0**	24	**1**	100	100
2006	Feb		**4**	100	**8**	27	100
2006	Mar		14	44	**3**	**1**	100
2006	**Apr**	**Blue Line First Ext.**	**1**	**0**	**2**	56	100
2006	May		**4**	**4**	**2**	57	100
2006	Jun		15	14	42	12	100
2006	Jul		69	69	71	74	74
2006	Aug		33	28	71	50	15
2006	Sep		40	**0**	**5**	**0**	**0**
2006	Oct		**0**	**0**	30	**1**	**1**
2006	**Nov**	**Blue Line Second Ext.**	**0**	**0**	**0**	**0**	10
2006	Dec		**0**	**0**	**8**	**0**	**1**

Shares less than 10 percent highlighted in bold

Table A2a Predicting Missing Observations at ITO: Entire Series, 2004-2006

	NO2		CO		PM2.5	
Share Missing in Entire Series (%)	18		14		42	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err**	**Coeff.**	**Std. Err.**
Red Line (second extension)	1.065***	0.202	0.571***	0.142	0.556***	0.162
Yellow Line (introduction)	0.662***	0.148	0.048	0.071	-0.541***	0.158
Yellow Line (first extension)	-0.358**	0.152	-0.295**	0.131	-0.434**	0.156
Blue Line (introduction)	-0.128	0.098	0.370**	0.130	-0.234**	0.110
Blue Line (first extension)	0.516***	0.132	-0.207	0.167	0.206*	0.123
Blue Line (second extension)	0.032	0.159	0.140	0.095	-0.037	0.177
Rainfall	0.019	0.020	0.011	0.016	0.025*	0.013
Relative Humidity	-0.003	0.011	0.012	0.008	0.012	0.010
Temperature	0.002	0.033	0.029	0.033	-0.050*	0.029
Wind Speed	-0.007*	0.004	-0.002	0.003	0.003	0.003
Workday	-0.031	0.027	0.019	0.023	0.006	0.024
Summer	-0.295**	0.103	-0.099	0.078	-0.252**	0.088
Winter	-0.037	0.085	0.008	0.053	-0.003	0.068
Diwali	0.019	0.085	0.052	0.052	0.025	0.104
Observations	25260		25260		25260	

The dependent variable is an indicator of whether the observation is missing. The explanatory variables are the same as in Table 4 with the addition of season fixed effects. Std. Errors are clustered at one week.
* indicates significantly different from zero at 10 percent level, ** at 5 percent level, and *** at 1 percent level.

Table A2b Predicting Missing Observations at Siri Fort: Entire Series, 2004-2006

	NO2		CO	
Share Missing in Entire Series (%)	27		52	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**
Red Line (second extension)	0.526**	0.231	-0.008	0.200
Yellow Line (introduction)	0.060	0.119	0.302**	0.138
Yellow Line (first extension)	-0.342**	0.172	0.082	0.137
Blue Line (introduction)	0.194	0.157	0.067	0.064
Blue Line (first extension)	0.222	0.235	-0.161	0.116
Blue Line (second extension)	0.382**	0.147	0.624***	0.152
Rainfall	0.010	0.017	0.036***	0.013
Relative Humidity	0.010	0.015	0.009	0.016
Temperature	-0.084**	0.035	0.024	0.032
Wind Speed	-0.008**	0.004	-0.005	0.003
Workday	0.026	0.030	0.033*	0.018
Summer	-0.186*	0.095	0.142	0.088
Winter	0.205**	0.078	0.095	0.084
Diwali	0.025	0.085	0.113**	0.051
Observations	25260		25260	

The dependent variable is an indicator of whether the observation is missing. The explanatory variables are the same as in Table 4, with the addition of season fixed effects Std. Errors are clustered at one week. * indicates significantly different from zero at 10 percent level, ** at 5 percent level, and *** at 1 percent level.

Table A3 Effect of Delhi Metro on NO2 at ITO: Entire Series, 2004-2006

	Coeff.	Std. Err.
Yellow Line First Extension (YL1E)	40.4	99.1
Blue Line Introduction (BLI)	-404.1**	161.6
t	-0.0002	0.0002
t squared	2.81E-08	3.61E-08
t cubed	-1.52E-12	1.78E-12
t*YL1E	-0.0085	0.0195
t squared*YL1E	5.71E-07	1.27E-06
t cubed*YL1E	-1.19E-11	2.72E-11
t*BLI	0.0623**	0.0278
t squared*BLI	-3.21E-06*	1.63E-06
t cubed*BLI	5.53E-11*	3.29E-11
Observations	20646	

The dependent variable is the natural logarithm of hourly CO at ITO. Control variables are the same as used in table 4, with the addition of season fixed effects and interactions between the discontinuity dummies and the time polynomial. Std. errors are clustered at one week. * indicates significantly different from zero at 10 percent level, ** at 5 percent level, and *** at 1 percent level.

Note: The coefficients on the discontinuity dummies are not directly comparable to the coefficients in tables 4 through 7, because here we have interacted the discontinuity dummies with the time polynomial.

Table A4 Annual Profits of the Delhi Metro Rail Corporation Ltd.

	From Traffic Operations	From All Operations[1]
	in 2004-05 prices (million rupees)	
2004-05	70.0	-7612.6
2005-06	285.0	8177.1
2006-07	843.4	2119.0
2007-08	984.7	1686.7
2008-09	1275.9	6122.6
2009-10	1640.7	-6163.7
2010-11	2805.0	278.1

1. These are before tax, not accounting for prior period adjustments and after accounting for depreciation and interest. Nominal figures were obtained from the Annual Reports of the Delhi Metro Rail Corporation Ltd. The CPI-IW was used to convert to real terms.

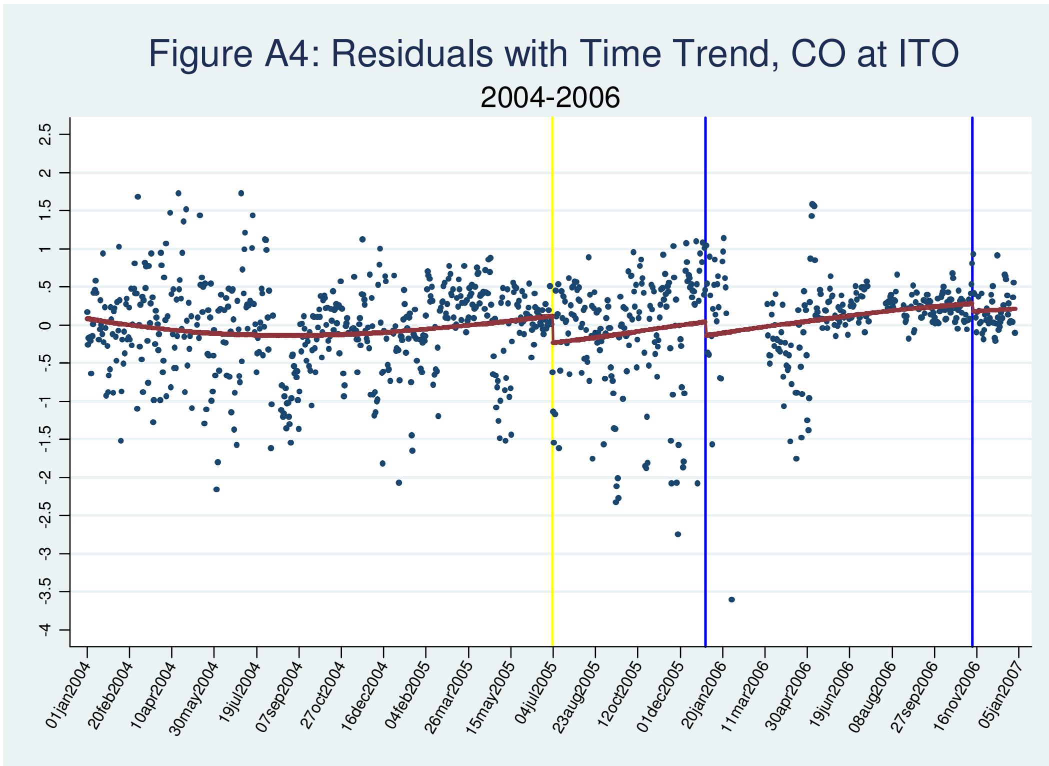


Figure A4: Residuals with Time Trend, CO at ITO
2004-2006
2.5
2
1.5
1
.5
0
-.5
-1
-1.5
-2
-2.5
-3
-3.5
-4
01jan2004
20feb2004
10apr2004
30may2004
19jul2004
07sep2004
27oct2004
16dec2004
04feb2005
26mar2005
15may2005
04jul2005
23aug2005
12oct2005
01dec2005
20jan2006
11mar2006
30apr2006
19jun2006
08aug2006
27sep2006
16nov2006
05jan2007

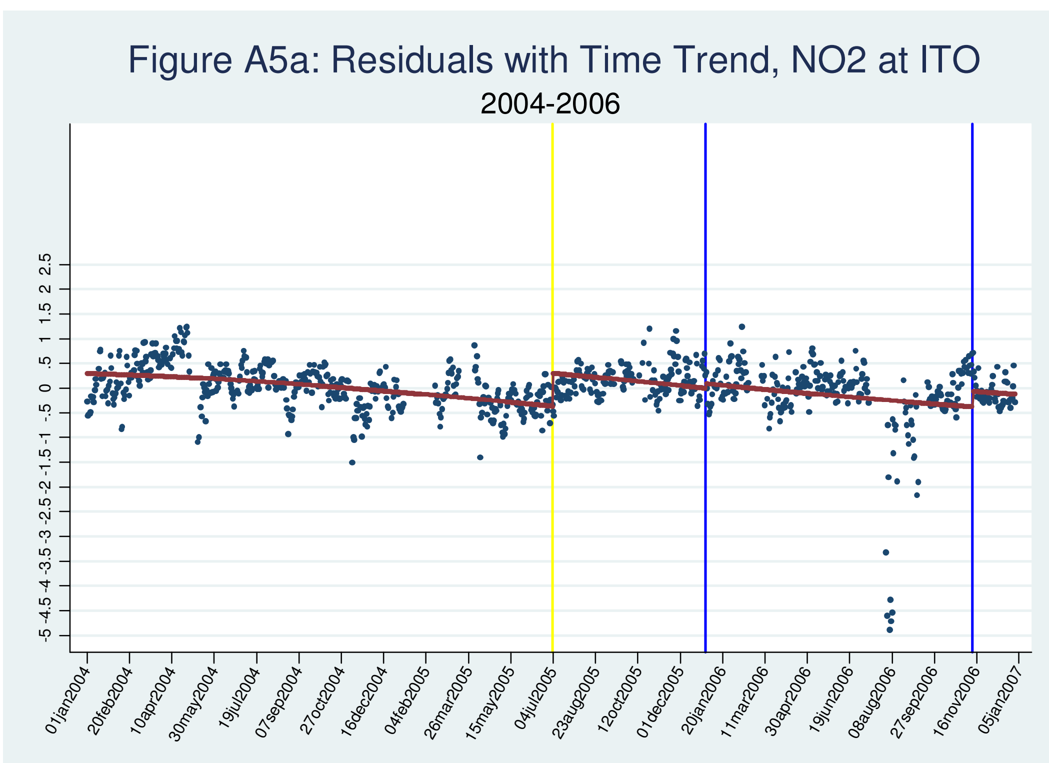
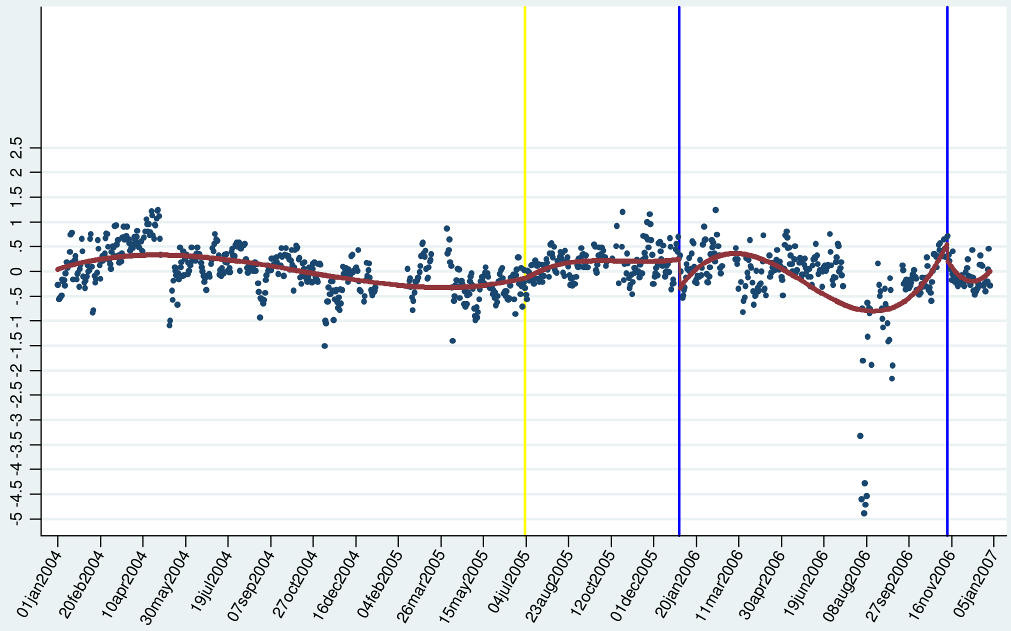


Figure A5a: Residuals with Time Trend, NO2 at ITO
2004-2006
2.5
2
1.5
1
.5
0
-.5
-1
-1.5
-2
-2.5
-3
-3.5
-4
-4.5
-5
01jan2004
20feb2004
10apr2004
30may2004
19jul2004
07sep2004
27oct2004
16dec2004
04feb2005
26mar2005
15may2005
04jul2005
23aug2005
12oct2005
01dec2005
20jan2006
11mar2006
30apr2006
19jun2006
08aug2006
27sep2006
16nov2006
05jan2007

Figure A5b: Residuals with Interacted Time Trend, NO2 at ITO

2004-2006



APPENDIX B

1 Calibration Exercise: Using independent inventory studies

We look at three pollution inventory studies for Delhi conducted in different years. The share of CO contributed by vehicles as reported in these studies is given in the table below.

Inventory Study	**Period for which share is applicable**	**Share of CO due to vehicles (in %)**
Gurjar (2004)	1990-2000	86
NEERI (2010)	2007	58
Guttikunda and Calorie (2012)	2010	28

Back calculating, if vehicles contributed 86% of CO at the start of our study period, and if CO drops by 34% (our impact estimate), then the new share of vehicles in CO pollution, assuming that the entire drop is coming from a reduction in vehicular emissions and the absolute contribution of other pollution sources remains unchanged, would be 79%. Compare this with 58% contribution reported by NEERI for 2007. If we repeat this exercise by assuming that vehicles contributed 58% of CO at the start of our study period, and again if CO drops by 34%, then the new share of vehicles (under the same assumptions above) would be 36%. Compare this with 28% contribution reported in the Guttikunda and Calorie for 2010.

Thus, our impact estimate of a 34% reduction would actually have to be larger in order to match the declining contribution of vehicles noted in the later inventory studies. Of course, our conclusion is based on the assumption that the absolute contribution of other pollution sources remains unchanged, and that the methodologies of these inventory studies are comparable.

2 Calibration Exercise: Comparing with Doll and Balaban (2013)

Doll and Balaban (2013) estimate reductions in several pollutants, including CO, for the year 2011 as a result of the Delhi Metro being available as an alternative mode of travel. Their methodology is heavily data dependent, and involves building a before and after scenario using data on various transport sector parameters such as total travel activity in Delhi (passenger kilometers travelled), total number of vehicles of various kinds (cars, buses, two-wheelers etc.), average distance travelled, vehicle occupancy, composition of pre-metro travel modes of DM ridership, fuel efficiencies, and fuel emission factors.

They estimate an annual DM ridership of 651 million passengers in 2011. They also estimate an average trip distance of 14.7 kilometers, which equates to 9.66 billion passenger kilometers. This they calculate to be

6.6 percent of Delhi's motorized travel demand. Using a primary survey they estimate that 44 percent of the DM ridership is from buses, 22 percent from cars, 25 percent from two wheelers, 5 percent from three wheelers and 4 percent from taxis. Combining this information with other model parameters such as fuel efficiencies and emission factors, they estimate that the DM resulted in a reduction of 6,545 tons of CO. Guttikunda and Calorie (2012) estimate the total CO emissions in Delhi to be 1.52 million tons in 2010. If we assume the same total emissions for 2011 as well, then we arrive at a 0.43 percent reduction in CO due to the DM. This is much lower compared to our estimate of 34 percent.

We refrain from commenting on this huge difference beyond making the following observations:

1) We identify the localized reduction in CO at a major traffic intersection (ITO in Central Delhi) over the three year period 2004-2006, while their estimate is for the whole of Delhi for the year 2011.

2) The two methodologies are very different and each has some limitations. While our RD identification strategy is not robust to the presence of sporadic and mobile sources of pollution, our data are actual measurements on pollution and weather obtained from monitoring stations located in Delhi. Their method relies on estimating travel sector parameters using data from multiple studies, and sometimes relying on estimates for other cities (e.g. their occupancy rates for cars is taken from average vehicle occupancy for Asian countries). Their method does not account for dynamic feedback effects such as improvement in car fuel efficiencies leading to greater use of cars. Their estimates would also change if actual fuel efficiencies and emission factors are different from what they use in their model.