

The Effect of Metro Expansions on Air Pollution in Delhi

*Deepti Goel
Sonam Gupta*



WORLD BANK GROUP

Development Economics Vice Presidency

Operations and Strategy Team

October 2015

Abstract

The Delhi Metro (DM) is a 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 for reductions in CO₂ emissions. Did the DM also lead to localized reduction in three transportation source pollutants? Looking at the period 2004–2006, 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. Results for NO₂ are also suggestive of a decline, while those for PM_{2.5} are inconclusive due to missing data. These impacts of pollutant reductions are for the short run. A complete accounting of all long run costs and benefits should be done before building capital intensive metro rail projects.

This paper is a product of the Operations and Strategy Team, Development Economics Vice Presidency. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at deepti@econdse.org and sgupta@impaqint.com.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

The Effect of Metro Expansions on Air Pollution in Delhi

Deepti Goel and Sonam Gupta

JEL codes: Q5, R4

Keywords: Air pollution, Delhi Metro, urban transport

Deepti Goel (corresponding author) is an assistant professor in the Department of Economics, Delhi School of Economics, University of Delhi, Delhi 110007; her email is deepti@econ.dse.org. Sonam Gupta is a research associate with IMPAQ International; her email is sgupta@impaqint.com. The authors are grateful to the journal's referees for their critical feedback that substantially improved the paper. They also 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 help in obtaining some of the data used in the paper. The authors are also grateful to Shantanu Khanna for his excellent research assistance.

The Delhi Metro (DM) is an electric-based mass rapid rail transit system mainly serving the Indian National Capital Territory (NCT) of Delhi. The NCT of Delhi covers an area of 1,483 square kilometers and has a population of 16.8 million people according to the Indian Census of 2011, making it one of the world's most densely populated cities.¹ The DM was introduced in 2002 and since then it is being continually extended within the NCT and adjoining areas. As of 2012, its total route length was 190 kilometers, and annual ridership was 0.7 billion (DMRC 2012).² 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 other studies which show that air pollution causes cardiovascular disease (Brook et al. 2010) and worsens asthma

1. According to a worldwide ranking of cities by City Mayors Statistics, Delhi ranked thirteenth in terms of population density with 11,050 persons per square kilometer. Mumbai ranked first, and Beijing ranked twelfth, with population densities of 29,650 and 11,500, respectively. The data are compiled from various sources and are the most recent available. Source url: <http://www.citymayors.com/statistics/largest-cities-density-125.html>. Last accessed on March 31, 2015.

2. As of 2014, the Beijing Subway had a route length of 527 kilometers and annual ridership of 3.4 billion. This information is accessed from Wikipedia. Source url: http://en.wikipedia.org/wiki/Beijing_Subway. Last accessed on March 31, 2015.

(Auerbach and Hernandez 2012). Turning to recent research in economics, Tanaka (2015) finds that regulations to curb pollution from coal-based power plants in China led to 3.29 fewer infant deaths per 1000 live births, amounting to a 20 percent reduction in infant mortality rate.³ 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. Arceo-Gomez et al. (2012) provide evidence from Mexico City that a 1 percent increase in particulate matter (PM_{10}) over a year led to a 0.42 percent increase in infant mortality, and the corresponding figure for CO is 0.23 percent. Currie and Walker (2011) find that exposure to vehicular emissions around toll plazas in the northeastern United States increased the likelihood of premature births and also resulted in low birth weight. Some other studies that document the adverse health consequences of air pollution for the United States 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.⁴ The Central Pollution Control Board (CPCB), the national authority responsible for monitoring and managing air

3. Greenstone and Hanna (2014) find that environmental regulations in India have been effective in reducing air pollution. However, in contrast to Tanaka (2009), they find an insignificant impact of reduced air pollution on infant mortality. For reasons discussed in their paper, they advise readers to be cautious when interpreting their result on infant mortality.

4. 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 $PM_{2.5}$ 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).

quality in India, finds that pollution in Delhi is positively associated with lung function deficits and with respiratory ailments (CPCB 2008a; CPCB 2008b). Guttikunda and Goel (2013) estimate that particulate matter present in Delhi in 2010 led to premature deaths ranging from 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.⁵ 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 the cases of NO_2 , CO , and $PM_{2.5}$, there were a large number of occurrences when their levels exceeded prescribed limits, while there were 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 were much lower at 3 and 0.1 percent, respectively.⁶ Given that SO_2 and O_3 were within permissible limits most of the time, our analysis focuses on NO_2 , CO , and $PM_{2.5}$.

5. Appendix table S1.1 in the supplementary appendix S1 presents the CPCB limits along with those prescribed by the WHO. WHO (2000) maintains that its limits are to be interpreted as guidelines, and individual countries may have different standards based on prevailing exposure levels, and social, economic, and cultural considerations. We measure violations vis-à-vis the CPCB limits, as they should be more relevant within the Indian context.

6. 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*.

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),⁷ and 86 percent of CO . In another study for Delhi conducted in 2007, NEERI (2010) reports that the contribution of vehicles 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.⁸ 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 .⁹ 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 nonlinear 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, 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,¹⁰ and the consequent vehicular emissions are a

7. NO_x refers to both nitrogen monoxide (NO) and nitrogen dioxide (NO_2).

8. The contribution of vehicles towards particulate matter reported in NEERI (2010) and in Guttikunda and Calorie (2012) includes the contribution of road dust as well.

9. Gurjar et al. (2004) do not report this figure for SO_2 .

10. Among the 44 reported million-plus cities in India, Delhi had the largest number of registered motor vehicles during 2011–12 with 7.4 million vehicles (TRW 2013). Goel et al. (2015) estimate that of all registered cars and

matter of serious concern.

Theoretical research from transport economics (Vickery 1969; Mohring 1972) postulates the existence of two counteracting effects on air pollution from introducing a new mode of public transportation. On the one hand, the 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 that 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.¹¹ This substitution away 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

two-wheelers in Delhi in 2011; 59 percent and 42 percent, respectively, are in-use. If we make a very conservative estimate of vehicles in use and assume that only 42 percent of all registered motor vehicles are in use, we get a figure of 3.1 million for in-use vehicles in Delhi. In 2013, the number of in-use vehicles in New York City was 2 million (Department of Motor Vehicles, New York State website; source url: <http://dmv.ny.gov/org/about-dmv/statistical-summaries>, last accessed on March 31, 2015). Thus, Delhi has at least 55 percent more in-use vehicles than New York City.

11. According to a report by the Delhi Metro Rail Corporation (DMRC, 2008), the DM has already taken the share of 40,000 vehicles.

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. As we explain below, due to the presence of sporadic sources of pollution in Delhi (such as spontaneous burning of waste), this approach is not ideal for short periods of analysis. We therefore rely on a three-year study period and argue that these sporadic sources of pollution cancel each other within this timeframe, resulting in reliable estimates.

Our analysis reveals that soon after some of the larger extensions of the DM there were significant reductions in at least some 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*. Additionally, there is suggestive evidence of 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 I we briefly describe the institution of the DM. The empirical strategy is explained in section II and the data sources are listed in section III. Section IV presents our empirical results and section V places them in the context of existing literature on air pollution. Section VI ends with policy recommendations.

I 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 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 2006.¹² Figure 2 presents a map of the DM rail network as of December 31, 2006. The map shows the different metro extensions, the two air pollution monitoring stations at *ITO* and *Siri Fort*, and the weather station at *Safdarjung*.

II EMPIRICAL STRATEGY

We use Regression Discontinuity (RD) to estimate the causal impact of the DM on pollution.¹³ 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

12. One reason for not studying the period before 2004 is that we do not have pollution data for it.

13. Lee and Lemieux (2010) provide an excellent exposition of this method.

identifying assumption.

Estimation Equation

We measure pollution using data from monitoring stations at two different locations within the city, *ITO* and *Siri Fort*.¹⁴ 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 are 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:

$$y_t^{l,m} = \theta_0^{l,m} + \theta_1^{l,m} DM_t^m + \theta_2^{l,m} \mathbf{x}_t + \theta_3^{l,m} \mathbf{P}(t) + u_t^{l,m} \quad (1)$$

$y_t^{l,m}$ 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

14. 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 kilometers 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.

each window it takes the value 1 for all time periods after the extension date, and 0 for periods before it.¹⁵ \mathbf{x}_t , is the vector of covariates and includes controls for weather;¹⁶ for hour of the day,¹⁷ day of the week, and interactions between these two; and for public holidays and festivals such as Diwali.¹⁸ $\mathbf{P}(t)$ 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).¹⁹ 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

15. 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.

16. 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.

17. Figures S1.1a through S1.1c in the supplementary appendix S1 show how the pollutants at *ITO* vary by hour of the day and by season. Once again we see that pollution is higher in winter compared to summer. Also, there is substantial intraday variation with peak levels reached between 8 pm and 2 am. Similar patterns are observed at *Siri Fort*. One plausible reason for why pollution peaks during these night hours could be a citywide ban on the entry and movement of heavy goods vehicles (mostly diesel powered trucks) between 6:00 am and 9:00 pm. The substantial intraday variation in pollution calls for inclusion of hour of the day fixed effects in order to improve the precision of our estimates. In most specifications season-fixed effects are not included because of short window lengths, typically nine weeks. Choice of window lengths is discussed later in this section.

18. 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.

19. Before Chen and Whalley, Davis (2008) used similar identification to estimate the effect of driving restrictions on air pollution in Mexico City.

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, a large part of which was due to joy rides.²⁰ These joy rides are expected to die out as the novelty of the metro fades away. 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 missing observations in each pollutant series. The last column of table 2 shows the share of missing observations.²¹ The best pollutant series is *CO* at *ITO*, for which 14 percent of observations are 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.²² 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

20. "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" (DMRC 2008).

21. We were informed by a CPCB official that missing data could be due to many reasons: power cuts, instrument failure, software malfunction when transferring data to storage device, and disruption in telephone. According to the official, none of these reasons are systematically linked to high or low pollution episodes. Later on, in section IV, we examine whether there is evidence for a systematic pattern to missing observations.

22. At the hourly frequency, the number of potential observations in a week is $168 = 24 \times 7$. The 20-percent rule implies that each week in our estimation window has at least $134 = 0.8 \times 168$ observations.

one week.²³

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.²⁴ One example is a citywide 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. 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.²⁵

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 gas (CNG). However, this happened in 2001, well 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 were 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.²⁶ We do not know of any other regulatory

23. 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.

24. An event that has a gradual effect on pollution will be captured by the time polynomial $\mathbf{P}(t)$ and therefore does not impede our analysis.

25. For them to be problematic, the discontinuous 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.

26. 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

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 fifteen to thirty 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 biofuel 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.

III 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

within Delhi is dampened.

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*.²⁷ 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.

IV RESULTS

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

Ridership Discontinuities

For each month, figure 3 shows the percentage change in average daily ridership on the DM over the previous month.²⁸ 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 shows a significant rise in average daily ridership for the month (or for the following month)²⁹ of each extension.

27. 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 renovation; and (d) *DCE* is in a mainly rural part of Delhi where sporadic burning of biomass and wood is widespread.

28. 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.

29. 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.

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 relatively lower population density in south-west Delhi where the extension took place.³⁰ 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.

Impact Estimates: Short Windows

In order to estimate equation (1) using fairly good quality data with fewer

30. 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).

missing values, table 3 shows the maximum window length (in weeks) around each extension on applying the at-most-20-percent-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.

For all extensions with at least nine weeks of good quality data,³¹ figures S1.2a through S1.2h in the supplementary appendix S1, try to visually examine whether there is a break in a pollutant series at the extension date using a nine-week window around the date.³² These plots suggest that there is a drop in pollutant level at the time of each extension, and in some cases this drop is large. Next, we estimate equation (1) to arrive at quantitative estimates.

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.³³ 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*

31. For $PM_{2.5}$ at *ITO* we only have a seven-week window of good quality data, and therefore we look at this shorter window of seven weeks for it.

32. The scatter points are daily averages of residuals obtained from a regression of (log) hourly pollutant on all the right-hand side variables in equation (1) except the extension discontinuity, DM_t^m , and the time polynomial, $\mathbf{P}(t)$. The overlaid curve depicts the fitted residuals from a regression of the scatter points on the extension discontinuity and a third order time polynomial.

33. 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 semi logarithmic equation.

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 to be noted 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 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 it could even explain the very large magnitudes for some of the estimates seen in tables 4 and 5.³⁴ As discussed in section II, these sporadic and mobile sources of pollution pose a threat to our identification strategy. 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 even if the exact magnitudes of reduction may not be those reported in tables 4 and 5.

Another explanation for the disappearance of effects in table 6 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-traveled. They reason that reduced congestion on roads, experienced soon after new roads are built, has a feedback effect that 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 incentivized the remaining drivers to drive more, and may have also added some new drivers, thus wiping out the initial

34. Additionally, it could also explain some of the inconsistencies in our impact estimates. As pointed out earlier in this section, we had expected the first extension of the Yellow Line to lead to a reduction in both NO_2 and CO .

However, while we see a substantial reduction in CO , the effect on NO_2 is not statistically significant (see tables 4 and 5). There is also a discrepancy across extensions in the sense that while the magnitude of effects on NO_2 and CO at *ITO* are comparable for the introduction of the Blue Line, for the first extension of the Yellow Line these are quite disproportionate (see table 5).

effects on pollution (as seen in table 6) and on road congestion. 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.

Robustness Checks

Here we present some robustness checks for the results seen in table 4. We also present results from a new specification that uses all the data for our study period, ignoring the fact that there are missing observations. For reasons stated below, this is our preferred specification.

Varying the Order of the Time Polynomial: Following CW, we have used a third order time polynomial. However, in specifications similar to ours, Davis (2008) uses a seventh order polynomial. In order to check that our main results are robust to the choice of polynomial order, we ran the regressions in table 4 using a fourth through seventh order polynomial and found similar estimates. The results for the fourth and the seventh order specifications are shown in tables S1.2a and S1.2b in the supplementary appendix S1.³⁵

Accounting for Persistence in Pollution: Next we examine whether controlling for lagged pollution alters the estimated impacts. We estimate the equation shown below (l, m superscripts have been suppressed):

$$y_t = \alpha_0 + \alpha_1 DM_t + \alpha_2 \mathbf{x}_t + \alpha_3 \mathbf{P}(t) + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \beta_4 y_{t-4} + e_t$$

The results are shown in tables S1.3a and S1.3b in the supplementary appendix S1. As expected, when we account for persistence, the instantaneous impacts are lower.³⁶ The two instances in table 4 where we had seen significant drops in

35. We also tried interacting the time polynomial with the discontinuity dummy, DM_t^m . We do not report these results as, in all cases, the discontinuity dummy drops out due to multicollinearity in our dataset.

36. The instantaneous impact is α_1 , and the cumulative effect is calculated by iteratively substituting for y in the

pollution continue to remain significant, although the magnitude of the cumulative effects is lower.

Artificial Discontinuities: For NO_2 and CO , we implemented the standard placebo test of using data only from the pretreatment period (before the extension is made) and introducing artificial discontinuities within this period. In some cases we did find significant effects. In the same vein, we estimated equation (1) using the introduction of the Yellow Line, which is characterized by an insignificant jump in ridership (see earlier discussion under “Ridership Discontinuities”). For this expansion, we had expected to find insignificant effects, but instead we found significant declines in both NO_2 and CO at *Siri Fort*. These perverse results could be due to sporadic and mobile sources of pollution in Delhi. While we admit that this is a potential threat to identification, the fact that we consistently get negative effects for all discontinuities in tables 4 and 5 and that, in most cases, these effects are larger when we reduce the time window, provides strong support that DM did reduce pollution. Below, we provide additional evidence to support this claim.

Effects on Nontransportation Source Pollutants: Our conjecture is that traffic diversion is the main mechanism through which the DM affects pollution in the short run. We can test this by looking at the effects of various extensions on the two nontransportation source pollutants, namely, sulphur dioxide, SO_2 , and ozone, O_3 . If our hypothesis is true, then, we should not find any effects on these pollutants. In fact, CW use this as a robustness check for their results. We first explain why using O_3 for such a robustness check is invalid in our context. Then, we implement the check using SO_2 and discuss the results.

Formation of O_3 crucially depends on the *ratio* of volatile organic

equation shown. The cumulative effect is then given by,

$$[1 + \beta_1 + \beta_1^2 + \beta_1^3 + \beta_1^4 + 3(\beta_1^2\beta_2) + 2\beta_1\beta_2 + 2\beta_1\beta_3 + \beta_2 + \beta_2^2 + \beta_3 + \beta_4] * \alpha_1.$$

compounds (*VOC*) to NO_x (Sillman, 1999). The relationship between O_3 , *VOC*, and NO_x is nonlinear, and given that there is evidence that NO_2 (and perhaps, *VOC*, as well), was impacted by the DM, ex-ante it is not possible to know whether or not the DM would affect O_3 , and if it does, in what direction. Hence, we do not use O_3 .

Table S1.4 in supplementary appendix S1 presents the results when we estimate equation (1) using SO_2 at *ITO* as the dependent variable. We examine impacts for the first extension of the Yellow Line and for the introduction of the Blue Line; the two discontinuities for which we found significant declines in transportation source pollutants in table 4.³⁷ We had expected to see no impact but table S1.4 shows large positive effects for both extensions. This may be explained by the presence of three coal based power plants within the National Capital Region of Delhi (marked as triangles in figure 2) during our study period. According to Gurjar et al., 2004, and Guttikunda and Calorie, 2012, 68 and 55 percent, respectively, of SO_2 in Delhi is generated by these power plants. Figure S1.3 in the supplementary appendix S1 shows wide variation in the monthly power production of these plants. More importantly, it shows that monthly output was either high or rising when the two extensions were made. Since SO_2 emissions are strongly correlated with power produced, there will be corresponding variations in SO_2 . Thus, it is possible that it is the coal plants that are behind the significant impacts seen in table S1.4. In any case, that the two metro expansions were accompanied by an increase in SO_2 increases the credibility that these expansions led to reductions in some transportations source pollutants (as seen in table 4).

Missing Data Patterns: Since our analysis so far has been based on examining particular segments of good quality pollutant data, it is important that

37. In this exercise using SO_2 , for both discontinuities, we use a nine week window and once again adopt the criteria of at most 20 percent missing observations for each included week.

observations should not be missing systematically. For each pollutant, table S1.5 in supplementary appendix S1, 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 in two ways.

First we examine whether there is any association between data being missing and the timing of metro expansions. To do this we use data for our entire study period, 2004–2006, and estimate equation (1) by replacing the outcome variable with an indicator variable for whether the pollutant measure is missing. In this specification we include all six discontinuity dummies (representing the six metro expansions undertaken in this period), and also include season fixed effects. Results are presented in tables S1.6a and S1.6b for *ITO* and *Siri Fort*, respectively, in supplementary appendix S1. At *ITO*, for NO_2 and $PM_{2.5}$ there is some evidence that observations are less likely to be missing in summer. However, given that for each pollutant some extensions are positive, some negative, and some insignificant, we conclude that the extensions by themselves are not systematically associated with the pollutant measure being missing.

Next we investigate whether observations are more likely to be missing on either high or low pollution days. This is done using the entire series for CO and $PM_{2.5}$ at *ITO*. The last column of table 2 shows that CO at *ITO* has 14 percent missing observations, while the corresponding figure for $PM_{2.5}$ is 42 percent. Given these shares of missing observations, we use a specification similar to equation (1), wherein missing status for $PM_{2.5}$ is regressed on whether or not CO is above the CPCB prescribed limit of $2000\mu g/m^3$ (proxy for high pollution). Other right hand side variables are the same except that we drop the discontinuity dummies and the time polynomial, and once again include season fixed effects. The results are shown in appendix table S1.7 in supplementary appendix S1. The coefficient on

CO being high is insignificant suggesting that missing status for $PM_{2.5}$ is not systematically associated with high pollution episodes.

Impact Estimates using Entire Series: In order to get correct magnitudes of impact estimates we still need to address the threat posed by sporadic and mobile sources of pollution in Delhi’s pollution inventory. To do this we examine the entire series (between 2004 and 2006) for CO at ITO which has the least share of missing observations (14 percent missing) and estimate the following equation:

$$y_t = \delta_0 + \delta_1 \mathbf{x}_t + \delta_2 \mathbf{P}(t) + \sum_{i=1}^M \gamma_i (DMi)_t + \varepsilon_t \quad (2)$$

The variables are similarly defined as in equation (1), and as above, \mathbf{x}_t now includes season fixed effects. Because we look at the entire series we include all extensions shown in table 3 for ITO and this is represented by the set of discontinuity dummies, $\{DMi\}$. We expect that the sporadic sources of pollution are likely to be evenly spread within a long window of three years and would therefore cancel each other. This is therefore our preferred specification.

Figure S1.4 in the supplementary appendix S1 (similarly constructed as figures S1.2a through S1.2h), 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.³⁸

We also carry out similar analysis for NO_2 at ITO , which has 18 percent

38. To check whether the result is robust to choice of polynomial order, we ran the regression using fourth through seventh order polynomial in time. The main result does not change: the first extension of the Yellow Line is the only one that is significant, with estimate values of -32.9 and -39.3 for the fourth and seventh order polynomial, respectively. We also estimate equation (2) using the entire series for SO_2 at ITO which has 16 percent missing observations. The only significant discontinuity was the introduction of the Blue Line for which the sign was positive. Thus, the result in table 7 is also robust to the check using nontransportation source pollutant SO_2 .

missing observations for the entire series. Figure S1.5a in supplementary appendix S1 (counterpart of figure S1.4 for CO), 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 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 S1.5b 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 . 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 table 8. 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. We note that the coefficient on the introduction of the Blue Line is negative and significant.³⁹

Evidence for Traffic Diversion: Table 9 shows the growth rates in registered private vehicles in Delhi and Mumbai over two five year periods, 1997 to 2002, and 2002 to 2007. In 1997–2002 neither city had a metro system. In the later period the DM was operational in Delhi, while Mumbai did not see any major change in its public transport infrastructure. For each city, we calculate the difference in growth rates between the two time periods. Compared to 1997–2002, Delhi showed a decrease in the growth rate of Cars and Jeeps in the post-metro period, while Mumbai showed an increase. For two-wheelers, both cities showed an increase in growth rates in 2002–2007 compared to the earlier period, but the

39. 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 .

increase was smaller for Delhi compared to Mumbai. This is suggestive of a traffic diversion effect being operative in Delhi after the metro was introduced.

V 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.

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 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 (our preferred specification) suggests that the effect is not transitory.

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 the introduction of the Blue Line 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*, is further away from the extensions considered here.

In supplementary appendix S2, we compare our estimate of a 34 percent reduction in *CO* with the one in Doll and Balaban, 2013. Doll and Balaban also study the effect of the DM on air pollution, but they use a different methodology to

do so. The comparison reveals that our impact estimate is many times higher than theirs. We refrain from commenting on the reasons for this difference beyond what we have stated in this appendix.

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.⁴⁰ 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.

Next, we compare the implied ridership-pollution elasticities across the two studies. 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. Why should DM

40. 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 an 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.

have a higher ridership-pollution elasticity compared to the TM? Note that the pre-metro (1995–96) *CO* level in Taipei was $1,030\mu\text{g}/\text{m}^3$, while the *CO* level for Delhi before the first extension of the Yellow Line was $2,212\mu\text{g}/\text{m}^3$.⁴¹ Given that baseline pollution in Delhi is about twice that of Taipei, a one percentage point change in pollution would translate into a much larger absolute reduction in pollution 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, compared to Taipei, a much larger proportion of Delhi's population uses public transport.⁴² For traffic diversion from private vehicles to take effect, Delhi would therefore need a much larger increase in ridership.

Benefits from Reduced CO Pollution

Drawing from literature that looks at health impacts of air pollution, we quantify the benefits of reduced *CO* levels in Delhi in terms of infant lives saved. Currie, Neidell, and Schmieder (2009) study the effect of pollution on infant health in New Jersey in the 1990s. They account for avoidance behavior, and estimate that a 1 part per million (ppm) reduction in *CO*, reduced infant mortality by 2.5 percent in proportion to the baseline risk. CW combine this with their estimate of a 15 percent reduction in *CO*, and conclude that the opening of the TM resulted in 1.7 infant lives being saved in the year following the metro opening. We undertake a similar exercise and find that the expansion of the DM (first extension of the Yellow Line), resulted in about 61 infant lives being saved in the year following

41. CW present pollution levels in parts per million (ppm). We converted their figures into micro grams per cubic meter ($\mu\text{g}/\text{m}^3$) 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.

42. 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).

the extension.⁴³ Viscusi (2008) argues that less developed countries tend to have a lower value of statistical life (VSL), consistent with their lower levels of income. He calculates a VSL of 1.2–1.5 million USD (2000 prices) for India, and 0.2–0.9 for Taiwan, as against 7 for the U.S. Using the lower bounds of VSL estimates for India and Taiwan the infant lives saved may be valued at 73.2 and 0.34 million USD (2000 prices), respectively. Given that baseline pollution levels in Delhi are higher than in Taipei, and perhaps, avoidance behavior is much less in Delhi due to lack of pollution related warnings, we conjecture that this figure is a lower bound for Delhi.

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 labor 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.

Supplementary appendix table S1.8 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

43. For Delhi in 2004, on average, there were 838 births per day and the infant mortality rate (IMR) was 13 per 1000 infant population (Govt. of NCT of Delhi. 2008b). The pre-extension CO level was $2212 \mu g/m^3 = 1.794ppm$. Given our impact estimate of 34 percent and a 2.5 percent IMR improvement estimate, annual infant lives saved is calculated as: $1.794 * 0.34 * 0.025 * 0.013 * (305870) = 61$. The corresponding estimate for Taipei (as shown in CW) is: $0.834 * 0.156 * 0.025 * 0.00666 * (77029) = 1.7$.

population and high rates of economic growth, and this should alleviate concerns around low ridership demand.⁴⁴ 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.⁴⁵ 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 are 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 buses are more appropriate where travel is more dispersed. It seems to us that Delhi meets the criteria that favor having a metro. For these reasons we are optimistic about the viability of the DM.⁴⁶

VI CONCLUDING REMARKS

We study the impact of the Delhi Metro (DM) on pollution in Delhi between 2004 and 2006. We find that soon after the first extension of the Yellow Line of

44. 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).

45. 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.

46. 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 and 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 labor; premiums for importing fuel; benefits accruing from reduced road congestion, accidents, and air pollution; and effects of redistribution of income among stakeholders. Their estimates are based on several assumptions regarding annual flows of costs and benefits during the entire lifetime of the DM. Evaluating the accuracy of these assumptions is beyond the scope of our study.

the DM, there was a significant reduction in carbon monoxide at a major traffic intersection in central Delhi. There is also suggestive evidence that the introduction of the Blue Line led to a decline in nitrogen dioxide at the same intersection. Although we find a favorable impact of the DM on Delhi's pollution soon after some of the larger metro expansions, 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.⁴⁷ Also, our impact estimates are for the short run, and quantifying the net social benefit of the metro network in the long run is beyond the scope of our analysis.

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 multipronged strategy needs to be adopted if Delhi wants to shed its distinction 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

47. The DM has a regenerative braking 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 tons of CO_2 because of its regenerative braking technology.

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 citywide parking policies to dissuade use of private vehicles, adopting uniform emissions standards throughout the country, constructing a bypass road around Delhi to eliminate interstate 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 US 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 more advanced economies or to pander to private firms who build and design these systems. We 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

Arceo-Gomez, E. O., R. Hanna, and P. Oliva. 2012. “Does the Effect of Pollution on Infant Mortality Differ between Developing and Developed Countries? Evidence from Mexico City.” National Bureau of Economic Research, Cambridge, MA; Working Paper 18349.

Auerbach, A., and M. L. Hernandez. “The Effect of Environmental Oxidative Stress on Airway Inflammation.” *Current Opinion in Allergy and Clinical Immunology* 2012; 12 (2): 133–39.

Block, M. L., A. Elder, R. L. Auten, S. D. Bilbo, H. Chen, J. C. Chen, D. A. Cory-Slechta, et al. “The Outdoor Air Pollution and Brain Health Workshop.” *NeuroToxicology* 2012; 33(5): 972–84.

Brook, R. D., S. Rajagopalan, A. C. Pope III, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, et al. 2010. “Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association.” *Circulation* 121: 2331–2378.

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

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

———. 2008b. “Study on Ambient Air Quality, Respiratory Symptoms and Lung Function of Children in Delhi.” Central Pollution Control Board, Delhi; Environmental Health Series 2.

———. 2012. “Ambient Air and Noise Pollution Levels - Deepawali 2012.” Central Pollution Control Board, Delhi; Press Release, November 13, 2012.

Currie, J., M. J. Neidell, and J. F. Schmieder. 2009. “Air Pollution and Infant

Health: Lessons from New Jersey” *Journal of Health Economics* 28 (3): 688–703.

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

Davis, L. W. 2008. “The Effect of Driving Restrictions on Air Quality in Mexico City.” *Journal of Political Economy* 116 (1): 38–81.

DMRC. 2008. “A Journey to Remember.” Public Relations Department, Delhi Metro Rail Corporation Limited, New Delhi.

———. 2012. “Annual Report 2011–12.” Delhi Metro Rail Corporation Limited, New Delhi.

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

Duranton, G., and M. A. Turner. 2011. “The Fundamental Law of Road Congestion: Evidence from US Cities.” *American Economic Review* 101 (6): 2616–52.

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

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

Govt. of NCT of Delhi. 2008a. “Economic Survey of Delhi 2008–2009.” Planning Department, Government of NCT of Delhi.

———. 2008b. “Statistical Abstract 2008.” Directorate of Economics and Statistics, Government of NCT of Delhi.

———. 2012. “Economic Survey of Delhi 2012–2013.” Planning

Department, Government of NCT of Delhi.

Greenstone, M., and R. Hanna. 2014. “Environmental Regulations, Air and Water Pollution, and Infant Mortality in India.” *American Economic Review* 104 (10): 3038–72.

Gurjar, B. R., J. A. Aardenne, J. Lelieveld, and M. Mohan. 2004. “Emission Estimates and Trends (1990–2000) for Megacity Delhi and Implications.” *Atmospheric Environment* 38 (33): 5663–81.

Guttikunda, S., and R. Goel. 2013. “Health Impacts of Particulate Pollution in a Megacity- Delhi, India.” *Environmental Development* 6 (1): 8–20.

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

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

Lee, D., and T. Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature* 48:281–355.

Litman, T. 2014. “Evaluating Rail Transit Criticism.” Victoria Transport Policy Institute, Victoria, BC.

Lleras-Muney, A. 2010. “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* 45 (3): 549–90.

Ministry of Urban Development. 2008. “Study on Traffic and Transportation Policies and Strategies in Urban Areas in India: Final Report.” Ministry of Urban Development, Delhi.

Moretti, E., and M. J. Neidell. 2011. “Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles.” *Journal of Human Resources* 46 (1): 154–75.

Mohring, H. 1972. "Optimization and Scale Economies in Urban Bus Transportation." *American Economic Review* 62 (4): 591–604.

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

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

O'Toole, R. 2010. "Defining Success: The Case Against Rail Transit." *Policy Analysis* 663:1–40.

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

Sreedharan, E. 2009. "Climate Change—An Opportunity for Sustainable Development: The DMRC Experience." *RITES Journal*: 81–88.

Tanaka, S. 2015. "Environmental Regulations on Air Pollution in China and their Impact on Infant Mortality." *Journal of Health Economics* 42: 90–103.

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

Vickery, W. 1969. "Congestion Theory and Transport Investment." *American Economic Review* 59 (2): 251–60.

Viscusi, W. K. 2008. "Value of Life." In S. N. Durlauf and L. E. Blume, eds., *The New Palgrave Dictionary of Economics. Second Edition*. London: Palgrave Macmillan.

WHO. 2000. "Air Quality Guidelines for Europe." World Health Organization, 2nd Edition, WHO Regional Publications, Copenhagen; European

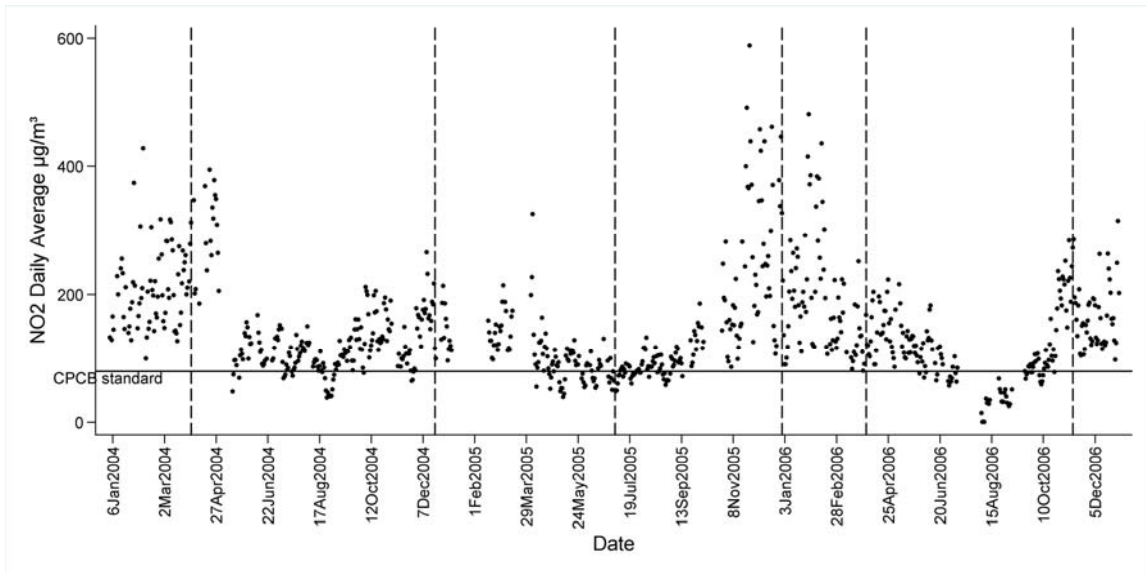
Series, No. 91.

———. 2006. “Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen dioxide and Sulphur dioxide. Global Update 2005.” World Health Organization, Regional Office, Copenhagen.

Winston, C., and V. Maheshri V. 2007. “On the Social Desirability of Urban Rail Transit Systems.” *Journal of Urban Economics* 62: 362–82.

FIGURES AND TABLES

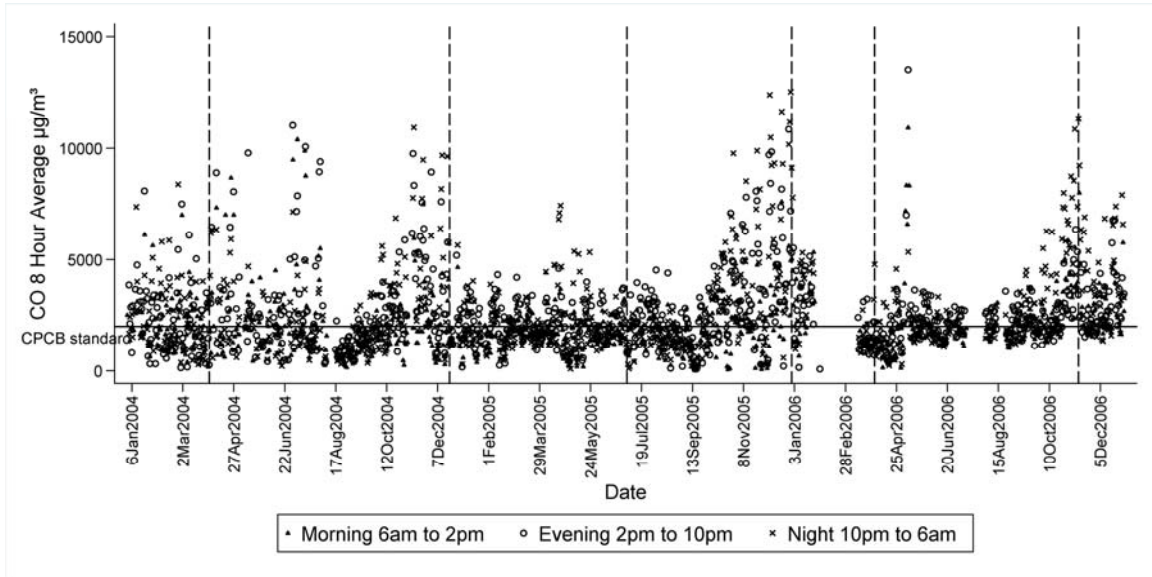
Figure 1A. Nitrogen Dioxide, ITO 2004–2006



Note: The dashed vertical lines indicate the dates of Delhi Metro expansion.

Source: Authors’ analysis based on pollutant data from the Central Pollution Control Board.

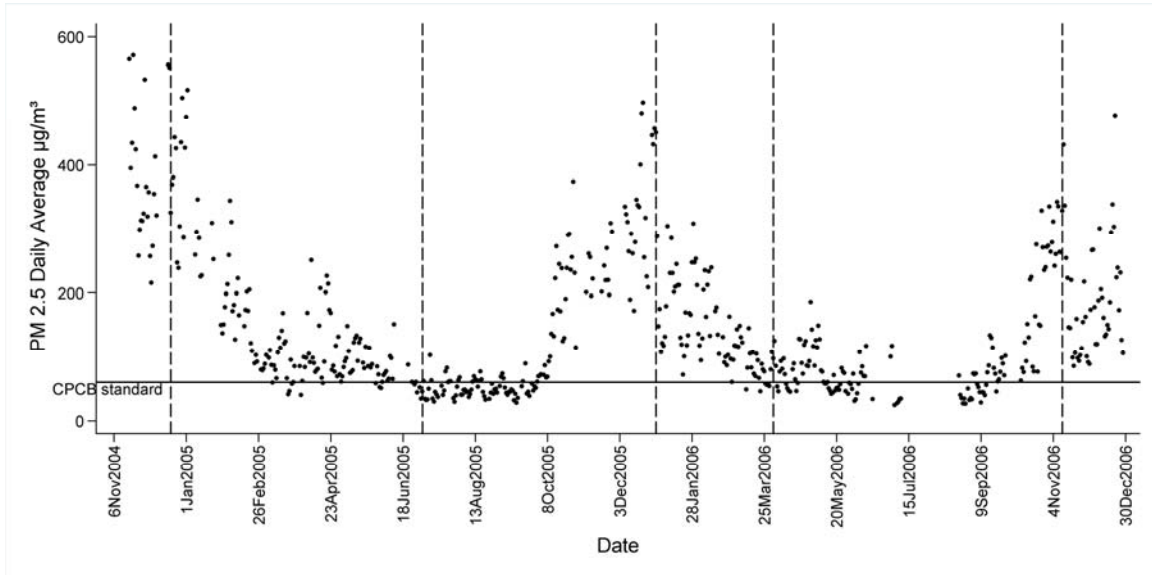
Figure 1B. Carbon Monoxide, ITO 2004–2006



Note: The dashed vertical lines indicate the dates of Delhi Metro expansion.

Source: Authors' analysis based on pollutant data from the Central Pollution Control Board.

Figure 1C. Particulate Matter 2.5, ITO 2004–2006

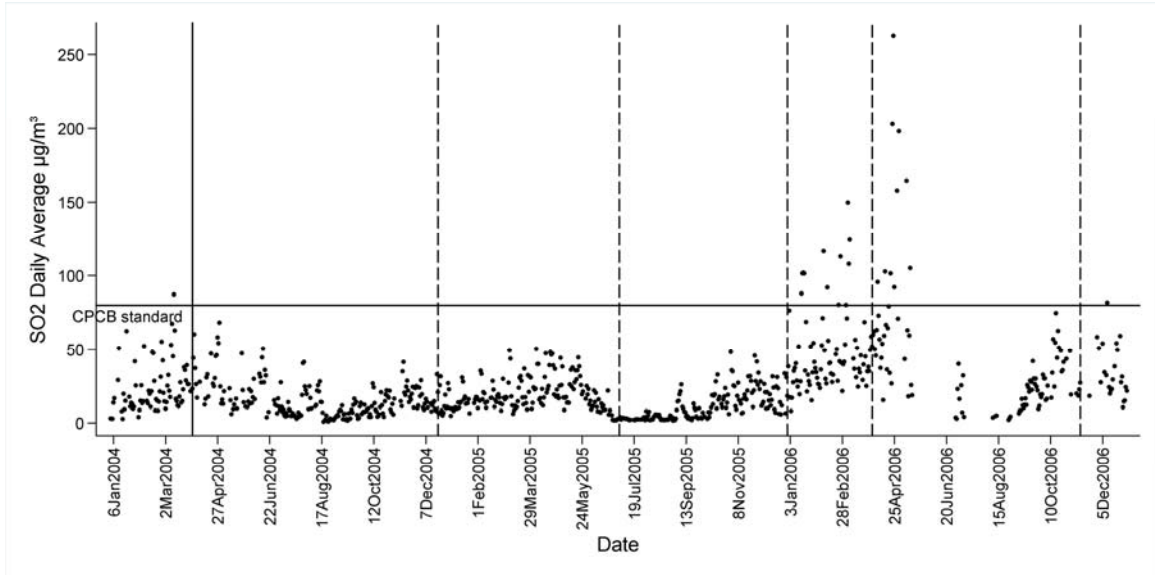


Notes: The dashed vertical lines indicate the dates of Delhi Metro expansion.

Data for PM 2.5 only available from November 2004 onwards.

Source: Authors' analysis based on pollutant data from the Central Pollution Control Board.

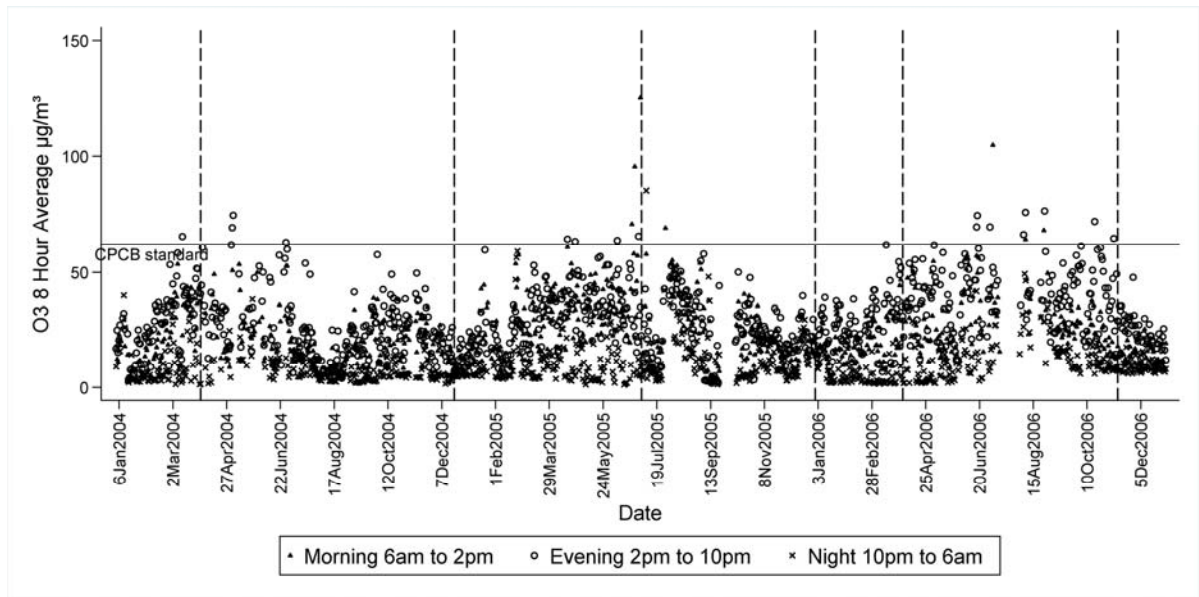
Figure 1D. Sulphur Dioxide, ITO 2004–2006



Note: The dashed vertical lines indicate the dates of Delhi Metro expansion.

Source: Authors' analysis based on pollutant data from the Central Pollution Control Board.

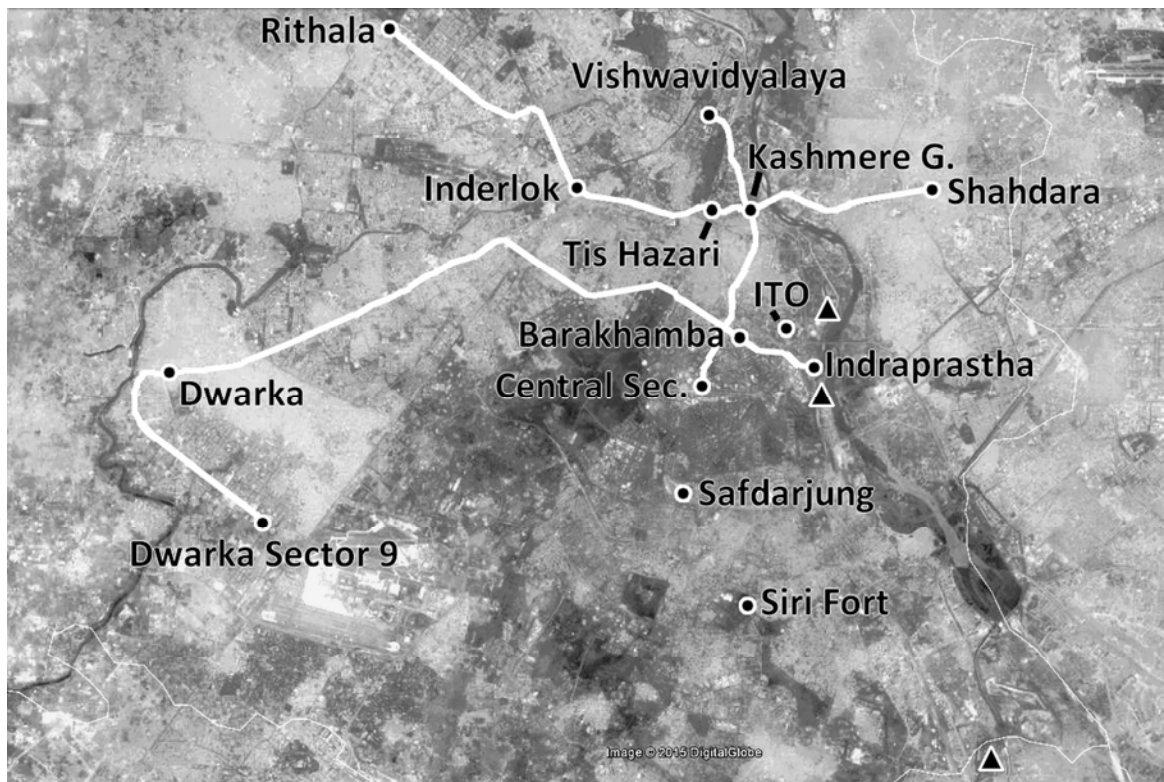
Figure 1E. Ozone, ITO 2004–2006



Note: The dashed vertical lines indicate the dates of Delhi Metro expansion.

Source: Authors’ analysis based on pollutant data from the Central Pollution Control Board.

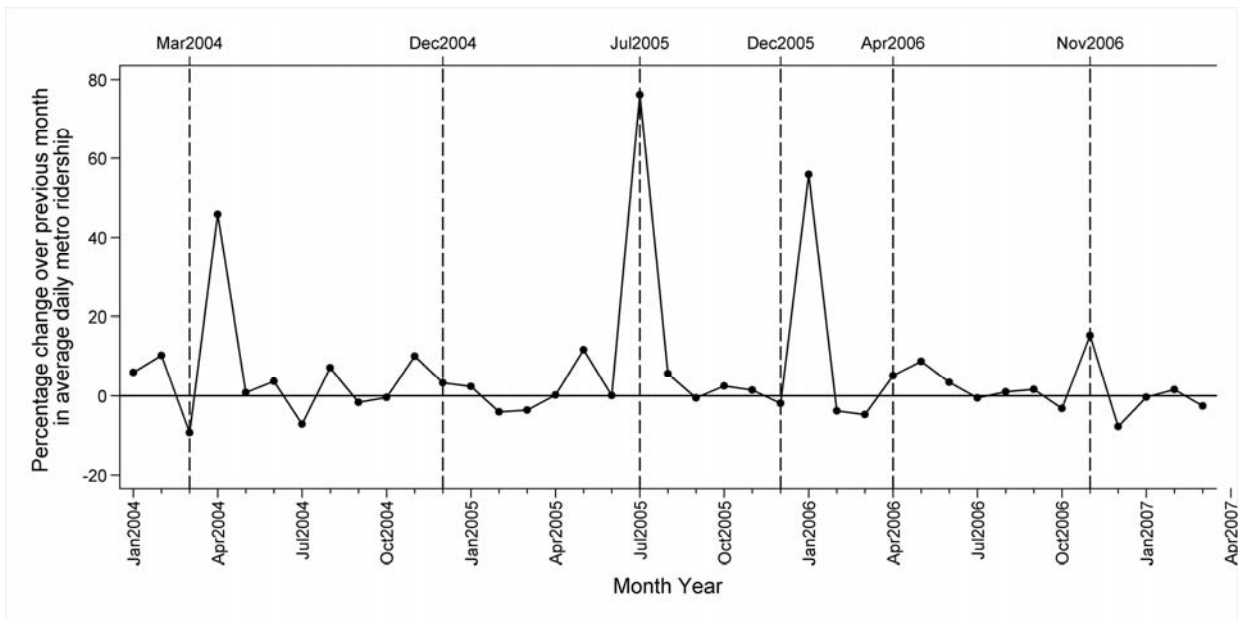
Figure 2. Delhi Metro Rail Network as of December 31, 2006



Notes: The two air pollution monitoring stations are at *ITO* and *Siri Fort*, and are approximately 9 kilometers apart. The weather station is at Safdarjung. The triangles represent the three coal based thermal power plants located within the National Capital Region of Delhi.

Source: Created by the authors using Google Earth and data from the Delhi Metro Rail Corporation Limited.

Figure 3. Change in Average Daily Metro Ridership over Previous Month, 2004–2006



Note: The dashed vertical lines indicate the dates of Delhi Metro expansion.

Source: Authors’ analysis based on ridership data from the Delhi Metro Rail Corporation Limited.

Table 1. Phase Wise Extension of the Delhi Metro from Inception to 2006

Extension	Segment (stations added)	Date of ext. dd-mm-yyyy	Average daily ridership (000s per day) ¹	% Change in average daily ridership ²
Extensions before our study period				
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				
Red Line (second extension)	Inderlok-Rithala (8)	31-03-2004	80.1	46
Yellow Line (introduction)	Vishwavidyalaya-Kashmere Gate (3)	20-12-2004	130.8	3
Yellow Line (first extension)	Kashmere Gate-Central Secretariat (6)	03-07-2005	143.2	76
Blue Line (introduction)	Barakhamba-Dwarka (21)	31-12-2005	270.4	56
Blue Line (first extension)	Dwarka- Dwarka Sector 9 (6)	01-04-2006	385.7	5
Blue Line (second extension)	Barakhamba-Indraprastha (3)	11-11-2006	451.0	15

Source: Delhi Metro Rail Corporation, DMRC

1. This refers to the average daily ridership for the month preceding the one in which the extension was introduced. However, for the two extensions carried out on the last day of the month (Red Line second extension and Blue Line introduction), it refers to the average daily ridership for the month in which the extension was introduced. For example, average daily ridership was 80.1 and 130.8 thousand per day in March 2004 and November 2004, respectively.
2. This refers to the month-to-month percentage change in average daily ridership for the month in which the extension was carried out. However for the two extensions carried out on the last day of the month, the change is for the month following the one in which the extension was carried out. For example, average percentage daily ridership in April 2004 was 46 percent higher than that in March 2004, while for December 2004 it was 3 percent higher than that in November 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)				(%)
NO ₂	145.5	109.3	0.226	1,140	18
CO	2,389.5	2,328.5	20.6	25,000	14
PM _{2.5} ²	144.6	131.8	1.69	1,020	42
	Siri Fort (in micrograms per cubic meter)				
NO ₂	45.0	52.5	0.771	805	27
CO	1,632.0	1,818.7	2.98	20,400	52
	Safdarjung				
Temperature	25.3	7.9	3.1	43.7	0.01

(deg. C)					
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' analysis based on data provided by the Central Pollution Control Board, 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' analysis based on 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	—

Source: Authors' analysis based on data described in the text.

1. Given the importance of analyzing PM_{2.5} 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

	NO ₂	CO	PM _{2.5}
	(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	—

Source: Authors' analysis based on data described in the text.

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	-3.1	—

31, 2005)		
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)

Source: Authors' analysis based on data described in the text.

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

Source: Authors' analysis based on data described in the text.

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.

Table 8. Effect of Delhi Metro on NO₂ at ITO: Entire Series, 2004–2006

	Coeff.	Std. err.
Yellow Line first extension	40.4	99.1
Blue Line introduction	-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
Number of observations	20646	

Source: Authors' analysis based on data described in the text.

The dependent variable is the natural logarithm of hourly NO₂ 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.

Table 9. Growth Rate in Registered Private Vehicles

	Cars and Jeeps		Two-wheelers	
	Delhi	Mumbai	Delhi	Mumbai
1997–2002 (a)	59.4	25.0	25.5	44.5
2002–2007 (b)	43.5	31.8	43.4	66.7
City difference (b) - (a)	-15.9	6.9	17.9	22.2
Delhi difference - Mumbai difference	-22.7	-4.3	—	—

Source: Various issues of Road Transport Yearbook, published by the Transport Research Wing, The Ministry of Road Transport and Highways.

Each figure in the table is a growth rate over a five-year period expressed in percent terms.

APPENDIX S1: SUPPLEMENTARY TABLES AND FIGURES

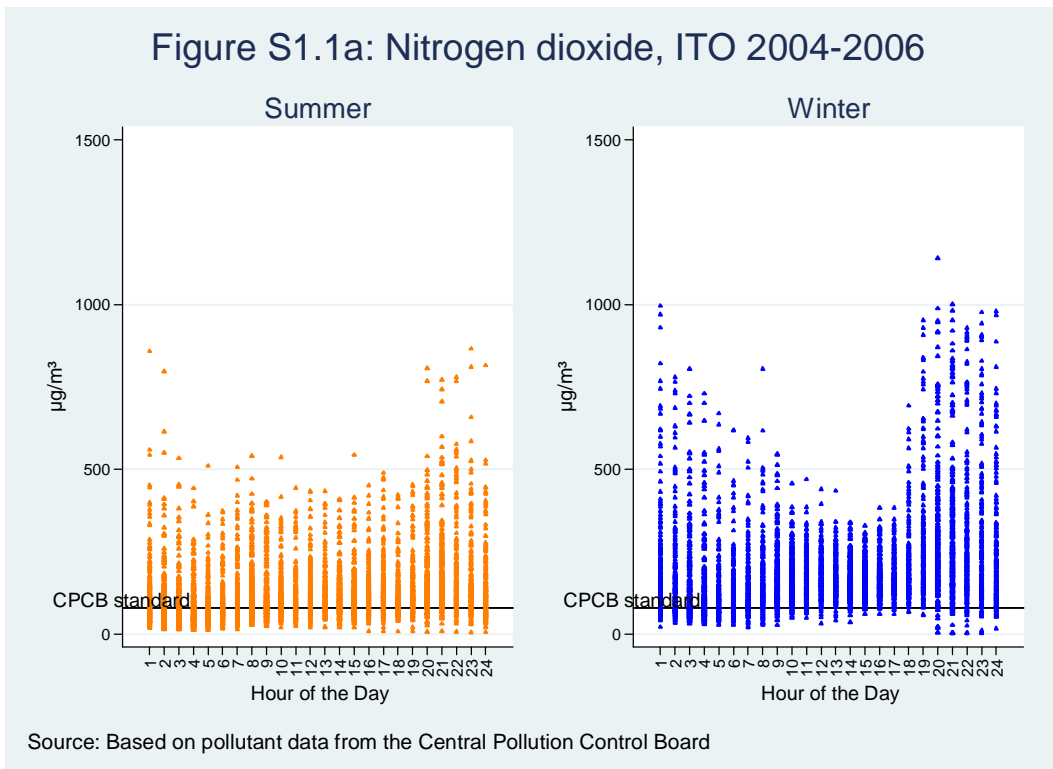
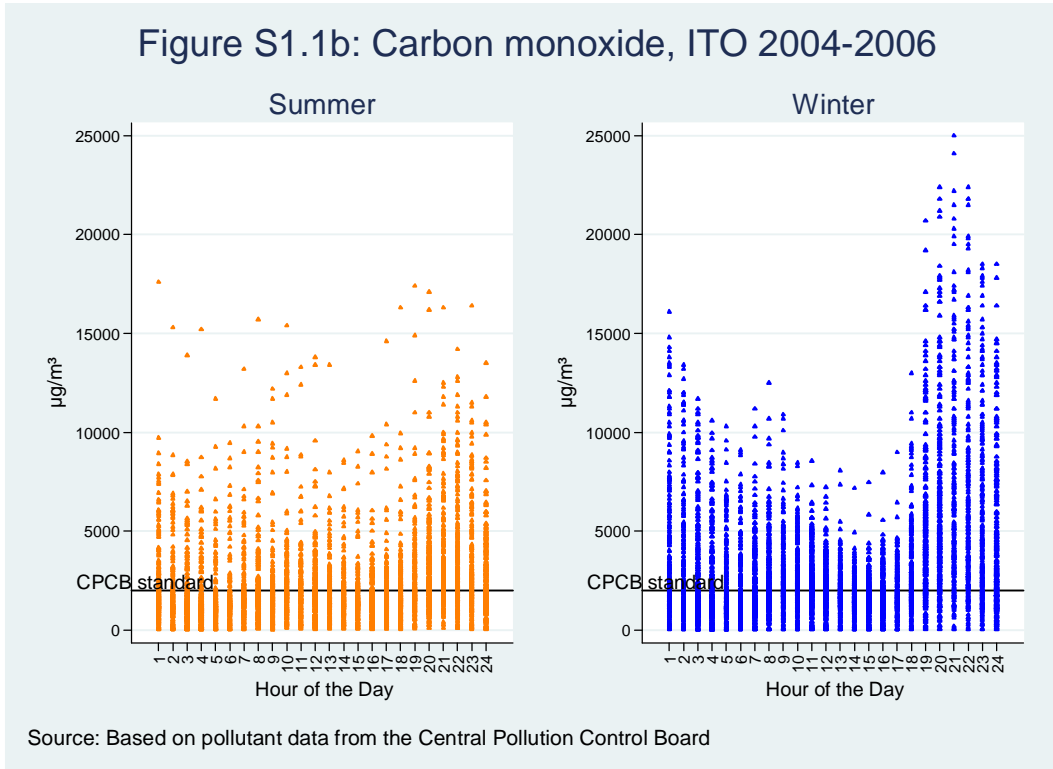
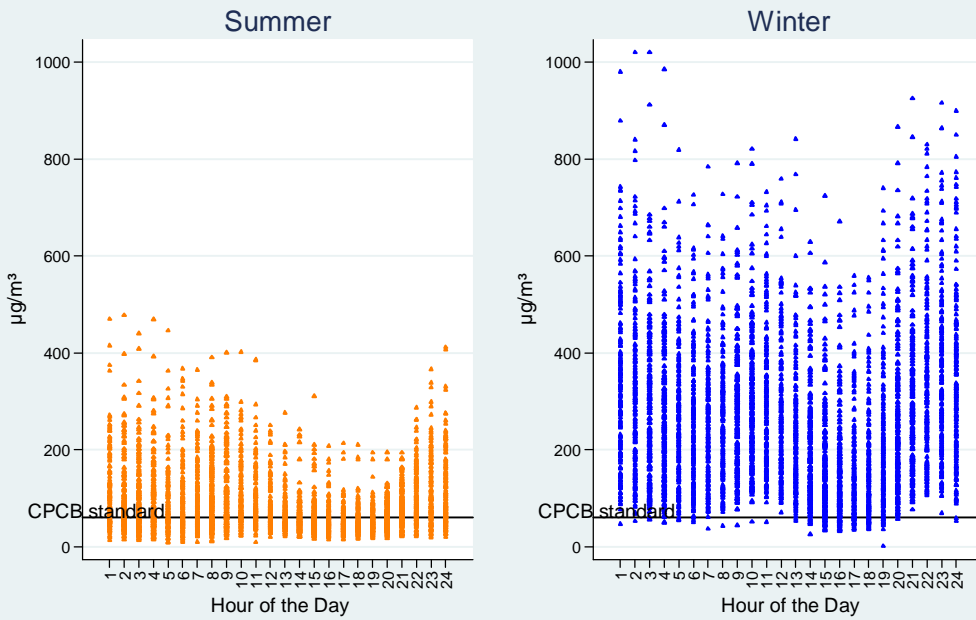
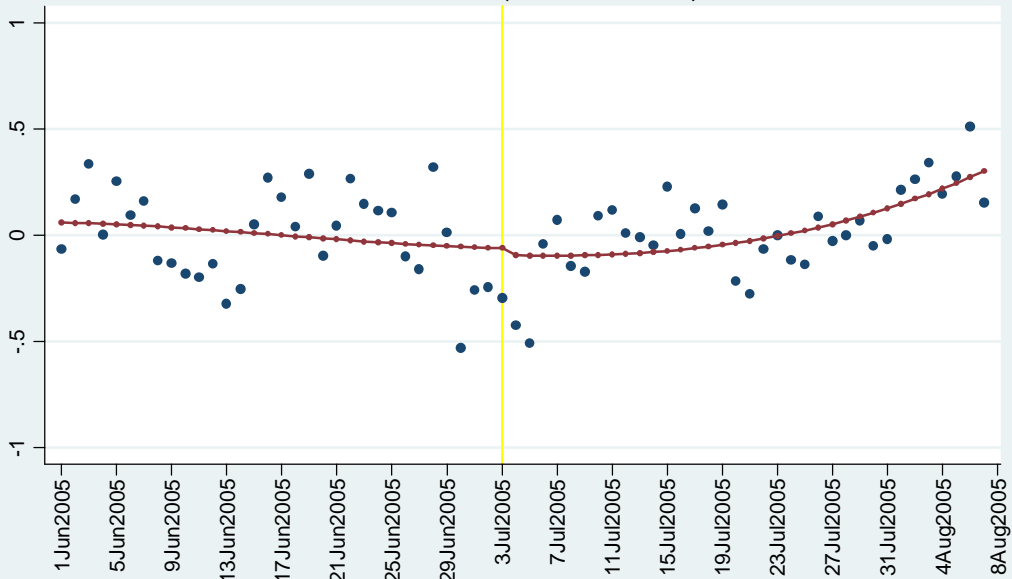


Figure S1.1c: Particulate Matter 2.5, ITO 2004-2006



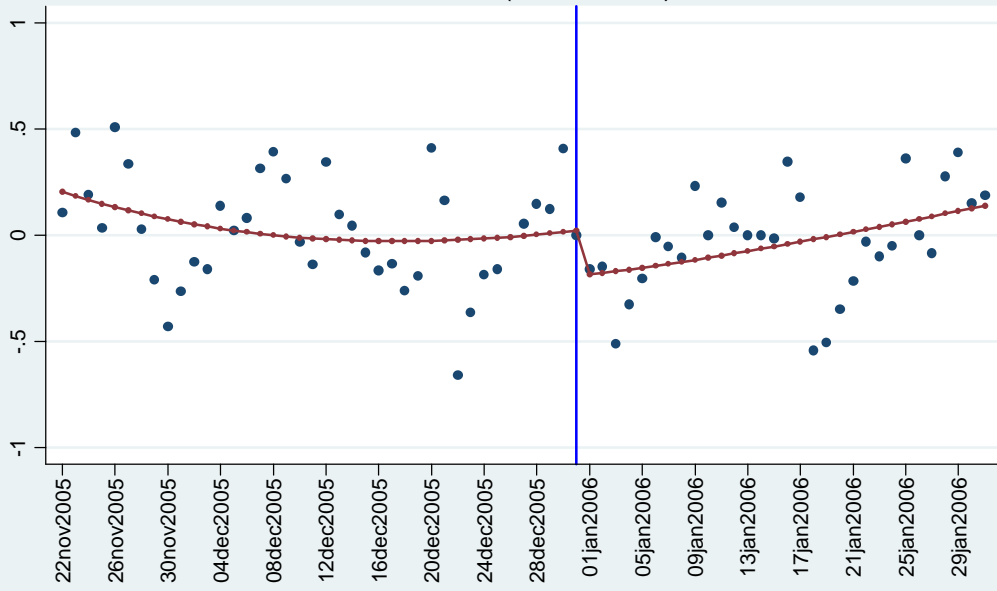
Source: Based on pollutant data from the Central Pollution Control Board

Figure S1.2a: Residuals with Time Trend, NO₂ at ITO
Yellow Line (first extension)



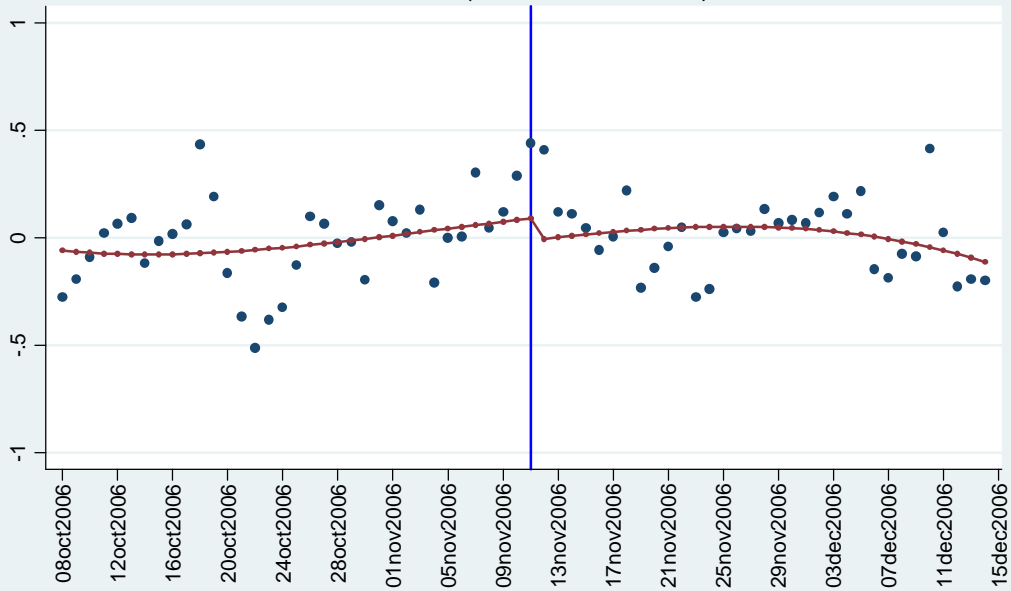
Source: Based on data described in the text

Figure S1.2b: Residuals with Time Trend, NO2 at ITO
Blue Line (introduction)



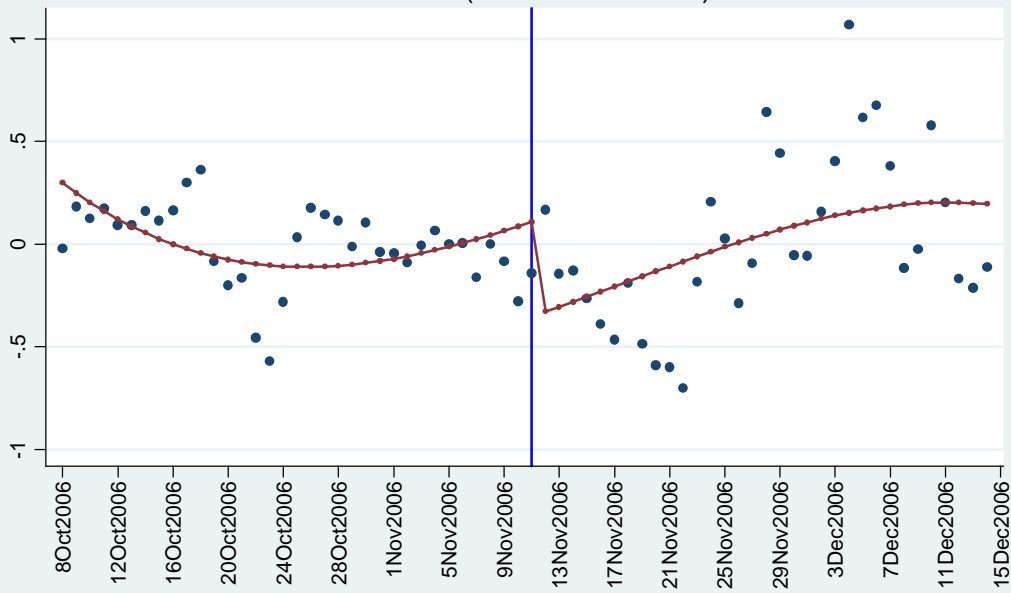
Source: Based on data described in the text

Figure S1.2c: Residuals with Time Trend, NO2 at ITO
Blue Line (second extension)



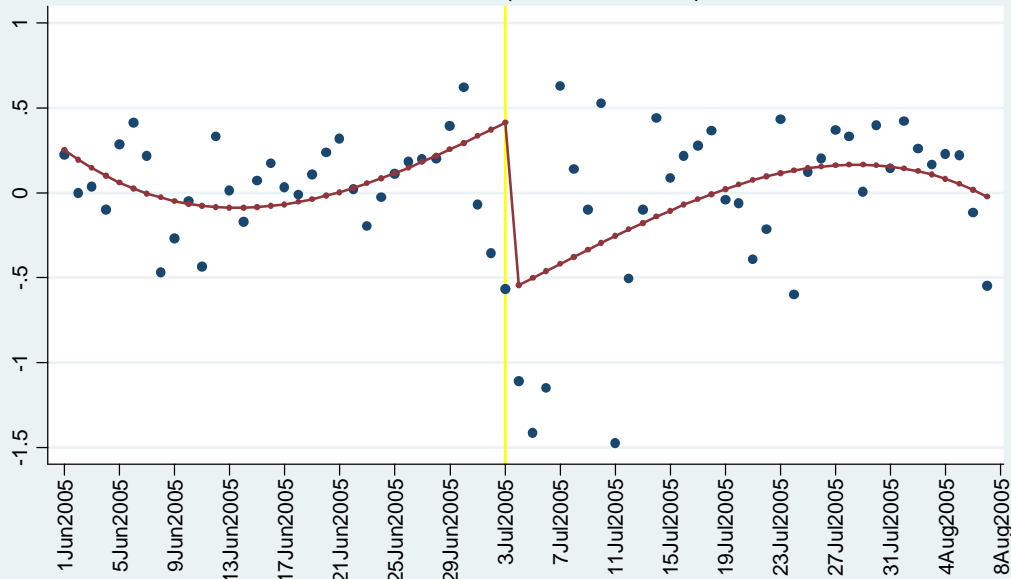
Source: Based on data described in the text

Figure S1.2d: Residuals with Time Trend, NO2 at Siri Fort
Blue Line (second extension)



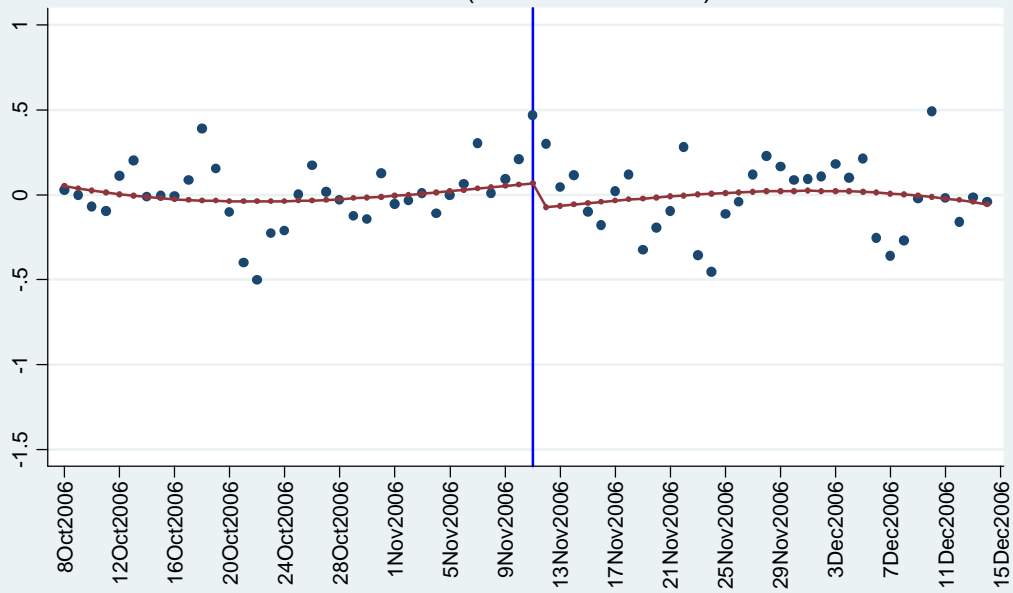
Source: Based on data described in the text

Figure S1.2e: Residuals with Time Trend, CO at ITO
Yellow Line (first extension)



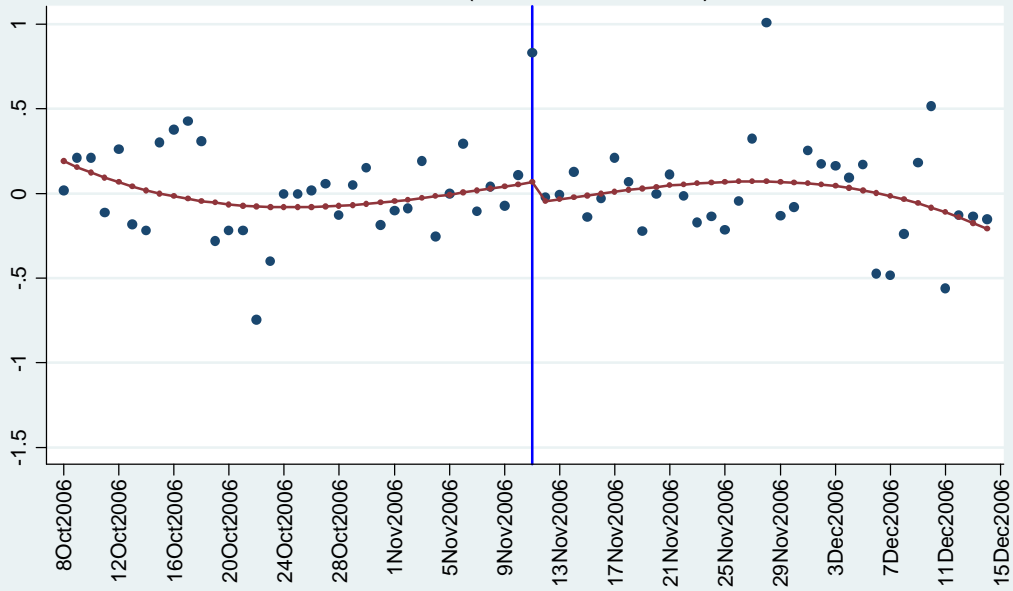
Source: Based on data described in the text

Figure S1.2f: Residuals with Time Trend, CO at ITO
Blue Line (second extension)



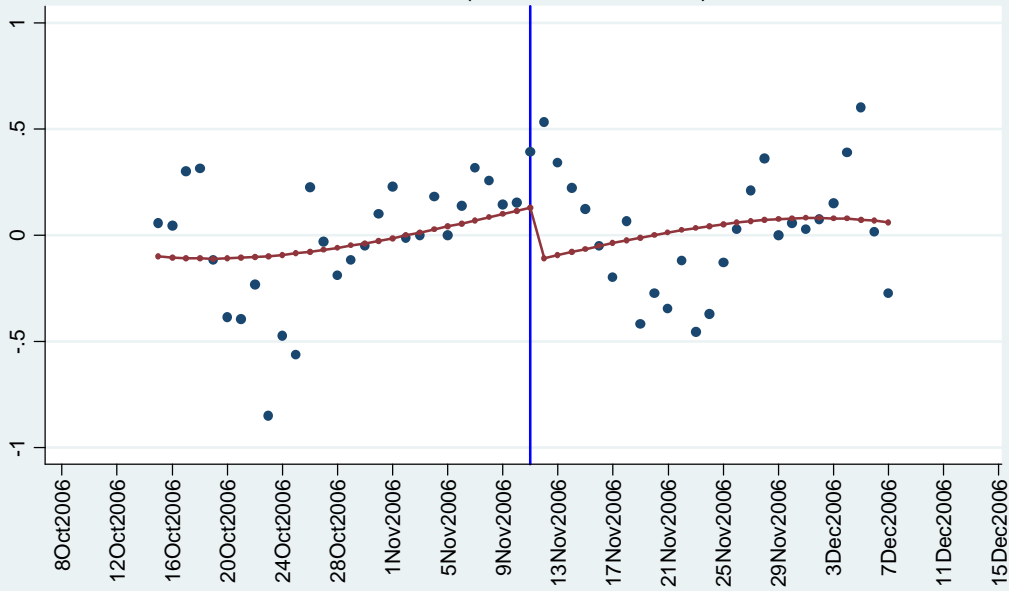
Source: Based on data described in the text

Figure S1.2g: Residuals with Time Trend, CO at Siri Fort
Blue Line (second extension)



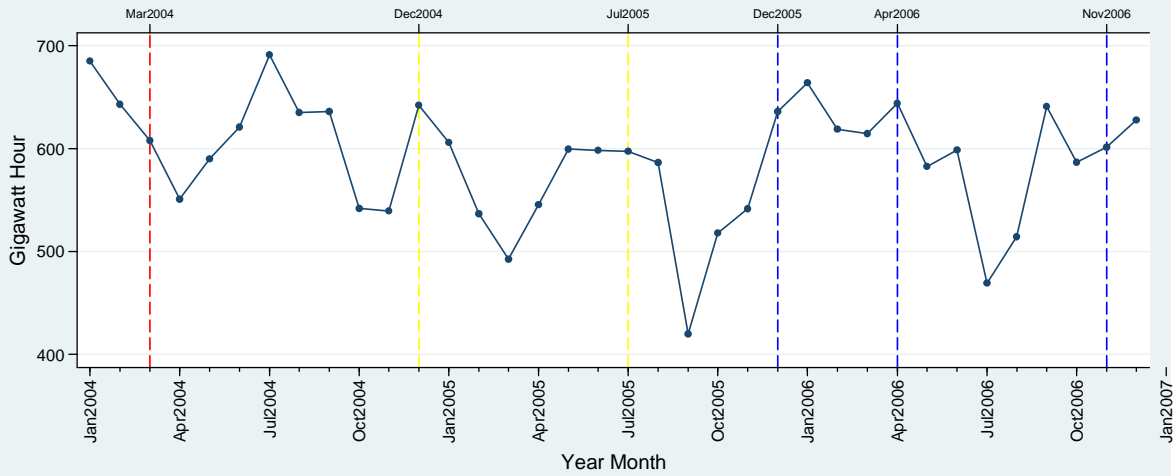
Source: Based on data described in the text

Figure S1.2h: Residuals with Time Trend, PM 2.5 at ITO
Blue Line (second extension)



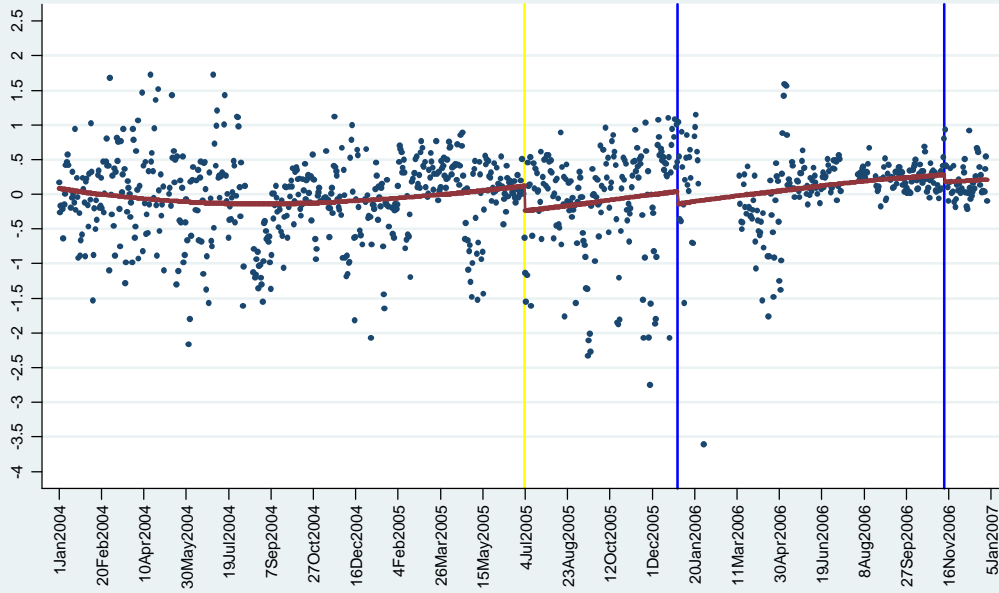
Source: Based on data described in the text

Figure S1.3: Monthwise Power Generation by Thermal Plants in Delhi, 2004-2006



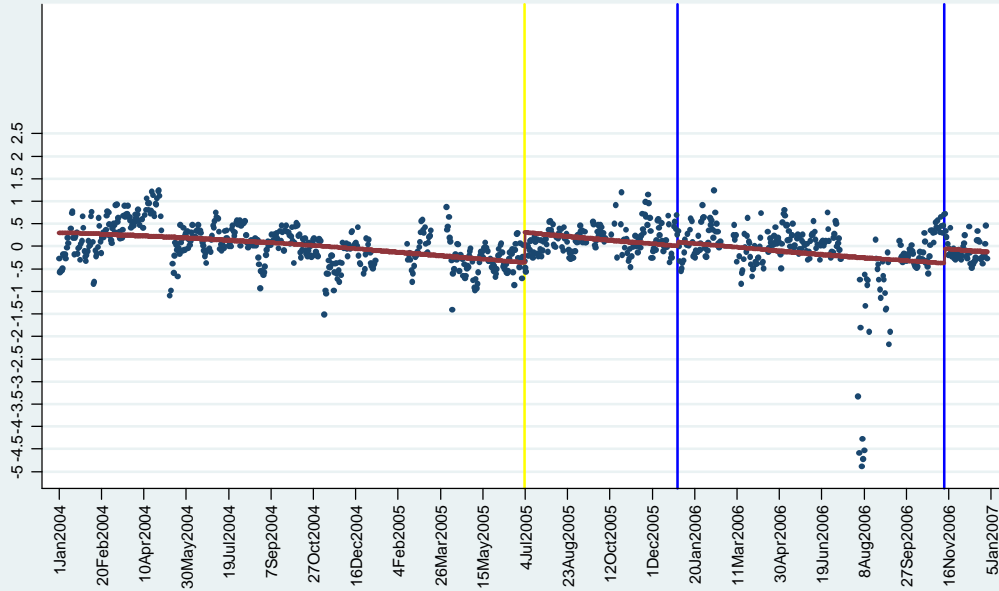
Note: Dashed vertical lines indicate Delhi Metro expansion episodes
Source: Based on data from the Central Electricity Authority

Figure S1.4: Residuals with Time Trend, CO at ITO
2004-2006



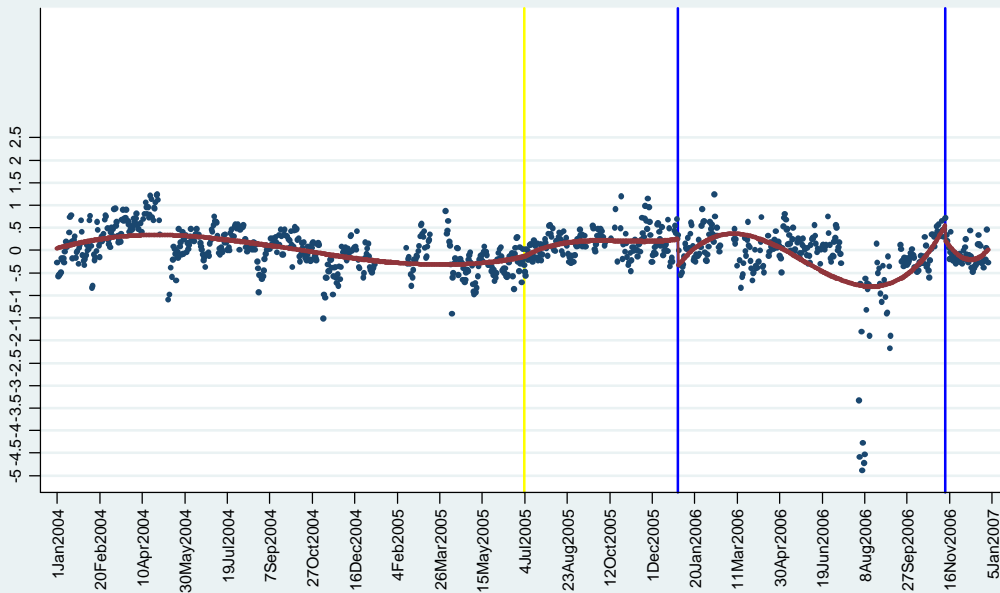
Source: Based on data described in the text

Figure S1.5a: Residuals with Time Trend, NO2 at ITO
2004-2006



Source: Based on data described in the text

Fig. S1.5b: Residuals Interacted w/ Time Trend, NO2 at ITO
2004-2006



Source: Based on data described in the text

Table S1.1 Air Quality Standards

Pollutant	Time Period	WHO Standard	CPCB Standard
		(in micro grams per cubic meter)	
NO₂	24 Hour Mean	Not Specified	80 ¹
CO	8 Hour Mean	10000	2000 ²
PM_{2.5}	24 Hour Mean	25 ³	60 ¹
SO₂	24 Hour Mean	20	80 ¹
O₃	8 Hour Mean	100	100

Source: CPCB 2011, WHO 2006 and WHO 2000

1. Allowed to exceed limit 2 percent of the time (7.3 days a year),
but not on consecutive days.

2. Allowed to exceed limit 2 percent of the time (22 8-hour blocks a year),

but not on consecutive days.

3. Allowed to exceed limit 3 days a year.

Table S1.2a Effect of Delhi Metro on Air Quality: Fourth Order Time Polynomial

Metro Extension	NO2	CO	PM2.5¹
	(Percentage change in level of pollutant)		
ITO			
Yellow Line First Extension (Jul 03, 2005)	-6.5	-69.6***	
Std. Error	(16.5)	(10.6)	
Observations	1457	1497	
Blue Line Introduction (Dec 31, 2005)	-30.5**		
Std. Error	(9.4)		
Observations	1639		
Blue Line Second Extension (Nov 11, 2006)	-10.6**	-13.1	-22.2
Std. Error	(9.9)	(8.3)	(19.6)
Observations	1605	1605	1268
Siri Fort			
Blue Line Second Extension (Nov 11, 2006)	-26.4	-3.0	
Std. Error	(23.9)	(9.9)	
Observations	1601	1532	

Source: Authors' analysis based on data described in the text.

1. Estimates for NO2 and CO pertain to a nine week window, and for PM 2.5 to a

seven week window.

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 S1.2b Effect of Delhi Metro on Air Quality: Seventh Order Time

Polynomial

Metro Extension	NO2	CO	PM2.5¹
	(Percentage change in level of pollutant)		
ITO			
Yellow Line First Extension (Jul 03, 2005)	-7.2	-67.6***	
Std. Error	(15.7)	(11.9)	
Observations	1457	1497	
Blue Line Introduction (Dec 31, 2005)	-28.8**		
Std. Error	(10.5)		
Observations	1639		
Blue Line Second Extension (Nov 11, 2006)	-10.8	-12.9	-21.7
Std. Error	(10.0)	(8.3)	(19.8)
Observations	1605	1605	1268
Siri Fort			
Blue Line Second Extension (Nov 11, 2006)	-27.7	-2.5	
Std. Error	(23.6)	(10.1)	
Observations	1601	1532	

Source: Authors' analysis based on data described in the text.

1. Estimates for NO₂ and CO pertain to a nine week window, and for PM 2.5 to a seven week window.

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 S1.3a Effect of Delhi Metro on Air Quality: Lagged Outcome Controls, ITO			
Metro Extension	NO₂	CO	PM_{2.5}
	(Percentage change in level of pollutant)		
Yellow Line First Extension (Jul 03, 2005)	-2.64	-47.81***	
Std. Error	(9.42)	(10.14)	
First Lag	0.53***	0.03***	
Std. Error	(0.05)	(0.004)	
Second Lag	0.15***	0.004**	
Std. Error	(0.04)	(0.002)	
Third Lag	0.11*	0.006**	
Std. Error	(0.06)	(0.002)	
Fourth Lag	0.02	0.004*	
Std. Error	(0.02)	(0.002)	
Cumulative Effect	-7.24	-50.21	
Observations	1457	1497	
Blue Line Introduction (Dec 31, 2005)	-11.58*		
Std. Error	(5.92)		
First Lag	0.23***		
Std. Error	(0.01)		

Second Lag	-0.03**		
Std. Error	(0.01)		
Third Lag	0.01		
Std. Error	(0.01)		
Fourth Lag	0.06***		
Std. Error	(0.01)		
Cumulative Effect	-15.47		
Observations	1639		
Blue Line Second Extension (Nov 11, 2006)	0.07	-8.65*	-0.28
Std. Error	(5.47)	(4.96)	(9.30)
First Lag	0.42***	0.02***	0.39***
Std. Error	(0.02)	(0.002)	0.05
Second Lag	-0.06**	-0.003*	-0.12***
Std. Error	(0.02)	(0.002)	(0.03)
Third Lag	0.01	0.001	0.08**
Std. Error	(0.03)	(0.001)	(0.04)
Fourth Lag	0.07***	0.003***	0.04
Std. Error	(0.02)	(0.001)	(0.03)
Cumulative Effect	0.12	-8.82	-0.45
Observations	1605	1605	1268

Source: Authors' analysis based on data described in the text.

Estimates for NO₂ and CO pertain to a nine week window, and for PM 2.5 to a seven week window. 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 S1.3b Effect of Delhi Metro on Air Quality: Lagged Outcome		
Controls, Siri Fort		
Metro Extension	NO2	CO
	(Percentage change in level of pollutant)	
Blue Line Second Extension (Nov 11, 2006)	-2.78	-1.83
Std. Error	(13.96)	(17.13)
First Lag	0.60***	0.03***
Std. Error	(0.09)	(0.004)
Second Lag	-0.22**	-0.01**
Std. Error	(0.08)	(0.002)
Third Lag	-0.01	0.004**
Std. Error	(0.04)	(0.001)
Fourth Lag	0.14***	0.001
Std. Error	(0.01)	(0.001)
Cumulative Effect	-5.37	-1.87
Observations	1601	1532
Source: Authors' analysis based on data described in the text.		
For both pollutants results pertain to a nine week window.		
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 S1.4 Robustness Check using Non-Transportation Source	
Pollutant	
Metro Extension	SO2 at ITO
	(Percentage change in level of pollutant)
Yellow Line First Extension (Jul 03, 2005)	89.7**
Std. Error	(33.9)
Observations	1469
Blue Line Introduction (Dec 31, 2005)	200.8**
Std. Error	(84.1)
Observations	1622
Source: Authors' analysis based on data described in the text.	
Results pertain to a nine week window.	
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 S1.5 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
Source: Authors' analysis based on data provided by CPCB.							
Shares less than 10 percent highlighted in bold							

	NO2		CO		PM2.5	
Share Missing in Entire Series	18		14		42	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
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

extension)						
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	

Source: Authors' analysis based on data described in the text.

The dependent variable is an indicator of whether the observation is missing (=1 if missing and =0 otherwise). 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 S1.6b Predicting Missing Observations at Siri Fort: Entire Series, 2004-2006				
	NO2		CO	
Share Missing in Entire Series	27		52	
	Coefficient	Std. Error	Coefficient	Std. Error
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	

Source: Authors' analysis based on data described in the text.

The dependent variable is an indicator of whether the observation is missing (=1 if missing and =0 otherwise). 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 S1.7 Predicting Missing PM2.5 using CO measure, at ITO: Entire Series, 2004-2006

	PM2.5 at ITO	
Share Missing in Entire Series	42	
	Coefficient	Std. Error
High_CO	-0.037	0.032

Rainfall	0.000	0.033
Relative Humidity	0.024	0.017
Temperature	0.007	0.053
Wind Speed	0.008	0.007
Workday	-0.020	0.029
Summer	-0.104	0.116
Winter	-0.155	0.118
Diwali	-0.018	0.151
Observations	22229	
Source: Authors' analysis based on data described in the text.		
The dependent variable is an indicator of whether PM2.5 observation is missing, (=1 if missing and =0 otherwise). The explanatory variables are the same as in Table 4, except that we drop all discontinuity dummies and the time polynomial, and include an indicator for whether CO measure is high (=1 if high and =0 otherwise) and also include 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 S1.8 Annual Profits of the Delhi Metro Rail Corporation Ltd.		
	From Traffic Operations	From All Operations
	Real Terms in 2004-05 prices (in 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
Source: Nominal figures were obtained from the Annual Reports of the		
Delhi Metro Rail Corporation Ltd. The CPI-IW, published by the Labour Bureau,		
Government of India was used to convert to real terms.		
1. These are before tax, not accounting for prior period		
adjustments and after accounting for depreciation and interest.		

APPENDIX S2: 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 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:

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

ii) 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.

iii) While our estimate is many times larger than Doll and Balaban's, it is comparable with the estimate in Chen and Whalley, 2012 who use similar identification strategy. Chen and Whalley find that the opening of the Taipei metro resulted in a 15 percent reduction in CO. We discuss their results in greater detail in section V.