NBER WORKING PAPER SERIES

SUBWAYS AND URBAN AIR POLLUTION

Nicolas Gendron-Carrier
Marco Gonzalez-Navarro
Stefano Polloni
Matthew A. Turner

Working Paper 24183
http://www.nber.org/papers/w24183

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2018

We are grateful to Tasnia Hussain, Fern Ramoutar, Mahdy Saddradini, Mohamed Salat, Farhan Yahya, and Guan Yi for assistance compiling the subway data. Assistance from Lynn Carlson and Yi Qi with the MODIS data and from Windsor Jarrod with computing is gratefully acknowledged, as are helpful conversations with Ken Chay and Michelle Marcus, and comments from Leah Brooks, Gabriel Kreindler, Remi Jedwab and many seminar participants. This paper is part of a Global Research Program on Spatial Development of Cities, funded by the Multi Donor Trust Fund on Sustainable Urbanization by the World Bank and supported by the UK Department for International Development. The project was made possible through financial support from SSHRC under grant #224995, the International Growth Center under grant #89337, the Connaught Fund, and the Ontario Work Study program. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Subways and Urban Air Pollution
Nicolas Gendron-Carrier, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A. Turner
NBER Working Paper No. 24183
January 2018
JEL No. L91,R11,R14,R4

ABSTRACT

We investigate the relationship between the opening of a city's subway network and its air quality. We find that particulate concentrations drop by 4% in a 10km radius disk surrounding a city center following a subway system opening. The effect is larger near the city center and persists over the longest time horizon that we can measure with our data, about eight years. We estimate that a new subway system provides an external mortality benefit of about $594m per year. Although available subway capital cost estimates are crude, the estimated external mortality effects represent a significant fraction of construction costs.

Nicolas Gendron-Carrier
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON, Canada M5S 3G7
nicolas.gendron.carrier@mail.utoronto.ca

Stefano Polloni
Department of Economics
Brown University
Box B
Providence, RI 02912
stefano_polloni@brown.edu

Marco Gonzalez-Navarro
Agricultural and Resource Economics
UC-Berkeley
207 Giannini Hall
Berkeley, CA 94720-3310
and University of California, Berkeley
marcog@berkeley.edu

Matthew A. Turner
Department of Economics
Brown University
Box B
Providence, RI 02912
and NBER
matthew_turner@brown.edu

1. Introduction

We investigate the effect of subway system openings on urban air pollution. We rely on two principal data sources. The first describes the universe of world subway systems. The second is a remotely sensed measure of particulates, Aerosol Optical Depth (AOD), recorded by the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua earth observing satellites between 2000 and 2014. These data allow us to measure airborne particulates everywhere in the world, monthly, with approximately 3km spatial resolution. Our strategy for establishing the causal effect of subways on AOD relies on a comparison of changes in particulates within a city around the time of a new subway system opening.

The data provides clear evidence of a structural break in an average city's AOD level around the time that it opens its subway network and does not indicate a trend break at any time in our sampling frame. The magnitude of this break is about 4% and is about constant over 48 post-subway months. In fact, the 4% decrease in AOD is evident over all 96 post subway months we observe, although estimates over this longer horizon are less precise and less well identified.

Consistent with the fact that subways tend to disproportionately serve central cities, we find that the effect of subway openings declines with distance from the city center. The data indicate that subways do not have the same effect on AOD in all cities. We find no evidence that richer, more populous, more rainy, or cities that open more extensive networks respond differently to subway openings, although we find suggestive evidence that more polluted and Asian cities experience larger AOD decreases following a subway opening. Finally, we find that subway openings have larger effects on AOD and on ridership than do the first expansions, and that the effect of the first expansion is larger than subsequent expansions.

Our findings are important for three reasons. Subways are often proposed as a policy response to urban air pollution. For example, Vollmer Associates *et al.* (2011) list air pollution reduction as an objective for New York City's 2nd avenue subway expansion. Our analysis provides a basis for assessing their cost effectiveness relative to other remediation policies. Apart from this paper, we are aware of only one study, Chen and Whalley (2012), that measures the effect of subways on air pollution. Like the present investigation, Chen and Whalley (2012) use an event study research design. Unlike the present study, Chen and Whalley (2012) study the opening of a single subway. In contrast, we study all of the 43 subway openings and 104 expansions that occurred anywhere in the world between February 2000 and December 2014. Thus, we dramatically improve on our ability to assess whether subway construction, in fact, reduces urban air pollution.

Second, our estimates of the reduction in pollution following subway openings and expansions, together with existing estimates of the health implications of particulates, allow us to calculate the value of averted mortality that follows from subway openings. We estimate that, for an average city in our sample, a subway opening prevents 9.4 infant and 221 total deaths per year. Using standard income-adjusted life values, this averted mortality is worth $21m and $594m per year,

respectively. These estimates do not include the effects of particulate reduction on morbidity or on productivity and so probably understate actual health benefits. Although available subway capital cost estimates are crude, the estimated external health effects represent a significant fraction of construction costs, particularly for subway systems with costs at the low end of the observed range.

Finally, little is known about transportation behavior in developing countries, and we shed indirect light on this important topic. First, and interestingly, we find no evidence that developing world and developed world cities respond differently to subways. This supports the idea that, at least in this regard, the two classes of cities are similar. Second, a back of the envelope calculation suggests that subways typically account for between 1.5%-10% of trips within a few years of their opening. Given what is known about the relationship between subway ridership and traffic, and between traffic and PM10, this level of ridership can plausibly account for the observed 4% reduction in particulates that follows a subway opening only if subways divert trips that would otherwise have occurred in particularly dirty vehicles or at particularly congested times. This is consistent with evidence that public transit serves the poor and that subways are much more heavily used at peak times for vehicle traffic.

2. Literature

While the effects of subways have been studied extensively, these studies have overwhelmingly focused on within-city variation in the relationship between proximity to a subway and housing prices or population density, e.g., Gibbons and Machin (2005) or Billings (2011). There are few studies which, like ours, exploit cross-city variation in subways. Among the exceptions, Baum-Snow and Kahn (2005) study a small sample of US subway systems to examine the relationship between subways and ridership, Voith (1997) also examines a cross-section of cities to investigate the relationship between subways and ridership. Finally, Gonzalez-Navarro and Turner (2016) exploits the same underlying panel data on subway stations that we use here to examine the effect of subway systems on long run urban population growth.

To our knowledge, the literature contains only a single paper (Chen and Whalley, 2012) examining the relationship between subways and urban air quality. Chen and Whalley (2012) examine changes in air pollution in central Taipei during the year before and after the opening of the Taipei subway in March of 1996. Chen and Whalley (2012) use hourly pollution measurements from several measuring stations in central Taipei, together with hourly ridership data over the same period. By examining the change in pollution levels around the time of the system opening, Chen and Whalley (2012) estimate an approximately 5%-15% reduction in Carbon Monoxide from the subway opening, about the same effect on Nitrous Oxides, but little effect on either Ozone or particulates.

It is useful to contrast this with our findings. We employ essentially the same research design, but consider the universe of subway cities over a considerably longer time horizon. On the other hand, we are restricted to a single measure of pollution, AOD, and to monthly frequencies. Our results are slightly different. Chen and Whalley (2012) find that the Taipei subway caused a 5-15% reduction in Carbon Monoxide. The confidence interval we estimate for AOD overlaps substantially with this range. However, Chen and Whalley (2012) find no effect on particulates, although our estimated 4% effect lies well within the confidence bounds of their estimated effect of the subway opening on Taipei's particulates. Together with the fact that our estimates indicate considerable heterogeneity in the effect of subways across cities, this suggests the difference between our finding and Chen and Whalley (2012) may reflect sampling error. Unfortunately, the opening of the Taipei subway predates the availability of the remotely sensed AOD data on which our analysis is based, so we cannot carry out a replication of the Chen and Whalley (2012) result in our sample.

3. Data

To investigate the effect of subways on urban air pollution we require data for a panel of cities describing subways, air pollution, and control variables. Our air pollution data are based on remotely sensed measures of suspended particulates. Our subways data are the result of primary data collection. We describe these data and their construction below, before turning to a description of control variables.

A *Subways*

We use the same subways data as Gonzalez-Navarro and Turner (2016) organized into a monthly panel. These data define a 'subway' as an electric powered urban rail system that is completely isolated from interactions with automobile traffic and pedestrians. This excludes most streetcars because they interact with vehicle and pedestrian traffic at stoplights and crossings. Underground streetcar segments are counted as subways. The data do not distinguish between surface, underground or aboveground subway lines as long as the exclusive right of way condition is satisfied. To focus on intra-urban subway transportation systems, the data exclude heavy rail commuter lines (which tend not to be electric powered). For the most part, these data describe public transit systems that would ordinarily be described as 'subways', e.g., the Paris metro and the New York city subway, and only such systems. As with any such definition, the inclusion or exclusion of particular marginal cases may be controversial.

On the basis of this definition, the data report the latitude, longitude and date of opening of every subway station in the world. We compiled these data manually between January 2012 and February 2014 using the following process. First, using online sources such as `http://www.urbanrail.net/` and links therein, together with links on wikipedia, we complied a list

of all subway stations worldwide. Next, for each station on our list, we record opening date, station name, line name, terminal station indicator, transfer station indicator, city and country. We obtain latitude and longitude for each station from GOOGLE maps. We use the subways data to construct a monthly panel describing the count of operational stations in each subway city between February 2000 and December 2014, the time period for which our air pollution data is available. By connecting stations with the most direct possible routes, we approximate network maps and can calculate route length.

171 cities had subways in 2014, 63 in Asia, 62 in Europe and 30 in North America. South America, Australia and Africa together account for the remaining 16. These 171 subway systems consist of 8,889 stations. Subway systems in Europe, North America and Asia are all about the same size, and those in South America are distinctly smaller. On average, the world stock of subway stations increased by about 200 per year and grew by about one third between 2000 and 2014.

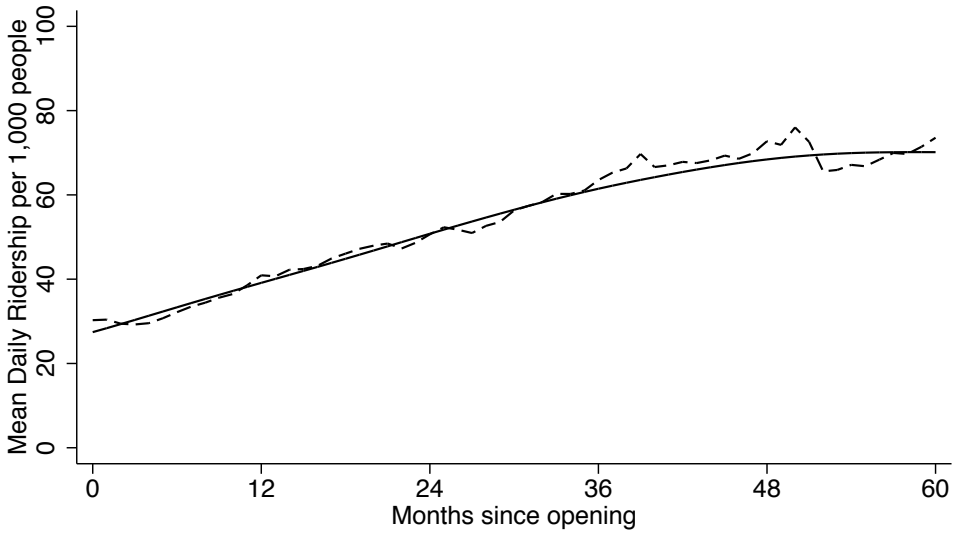
Our data on subway systems begins in the 19th century. However, our satellite pollution measures are more recent. Thus, our analysis will rely on subway openings that occurred between 2000 and 2014. Table A.2 lists all 43 subway system openings in the world between 2000 and 2014, by date, together with basic information about the cities where they are located. Subways opened fairly uniformly throughout the period and the average opening date is February 2007. During 2000-2014, an average system at the end of its first year of operation had a route of length 19.2km and 14.3 stations, usually on one line, but sometimes on two or more.

Our analysis hinges on the ability to observe a subway city for some time before and after an opening. Thus, we face a trade-off between sample size, the length of time we observe cities, and maintaining a constant sample of cities. While we experiment with other study windows, our primary econometric exercise considers the change in AOD in the period extending from 18 months before until 18 months after a subway opening. Since the AOD data cover February 2000 to December 2014, to base this exercise on a constant sample of cities, we must restrict our attention to subways that open between August 2001 and July 2013. In Table A.2, we see that these are the 39 cities beginning with Rennes and ending with Brescia. To consider the effect of subways over a longer horizon, in some specifications we consider a longer period after the subway opening. In this case, maintaining a constant sample of cities for each month requires that we drop cities that have openings near the end of the study period.

In theory, our research design could confound the effects of system openings with those of expansions that occur soon afterward. In fact, this is rare in our data. Among the 39 cities that make up our main sample, only 4 experienced an expansion of its subway system within 18 months of the system opening. We also experiment with dropping cities that experience an expansion soon after an opening from our sample. Our results are robust to such changes in sample.

For 30 of these 43 cities we are able to gather ridership data describing unlinked trips, mostly

Figure 1: Daily ridership per capita



Note: Graph depicts average daily passengers on subway per 1,000 people in metropolitan area, as well as a locally weighted regression of the series.

from annual reports or statistical agencies.[1] Ridership is reported at the monthly level for 19 of these 30. For the other 11, we interpolate to calculate monthly ridership from quarterly or yearly data. Table A.2 also reports mean daily ridership for each city where data is available, at the end of the first year of the system's operation. For an average city in our sample, about 97,514 people rode the subway on an average day in the twelfth month of the system's operation.

Figure 1 shows the evolution of ridership as a function of time from opening for the first five years of system operation. The horizontal axis in this figure is months from the opening date. The vertical axis is mean daily ridership per 1000 of city population. We see that ridership doubles over the three years following system opening and stabilizes at about 60 riders per thousand residents.

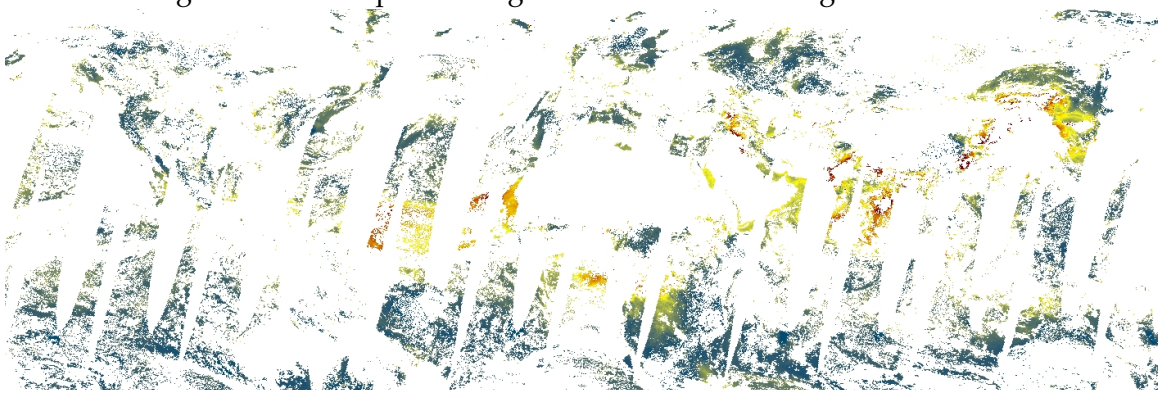
Finally, we determine the date when construction began for each subway opening in our sample. On average, construction begins 80 months prior to opening. The 25th and 75th percentile of construction duration are, respectively, 46 and 97 months. As we describe below, this variation in construction duration is important because it allows us to check whether the effect of subways on pollution is determined by the opening or the construction of the system.

B *Aerosol Optical Depth measurements from the Terra and Aqua earth observing satellites*

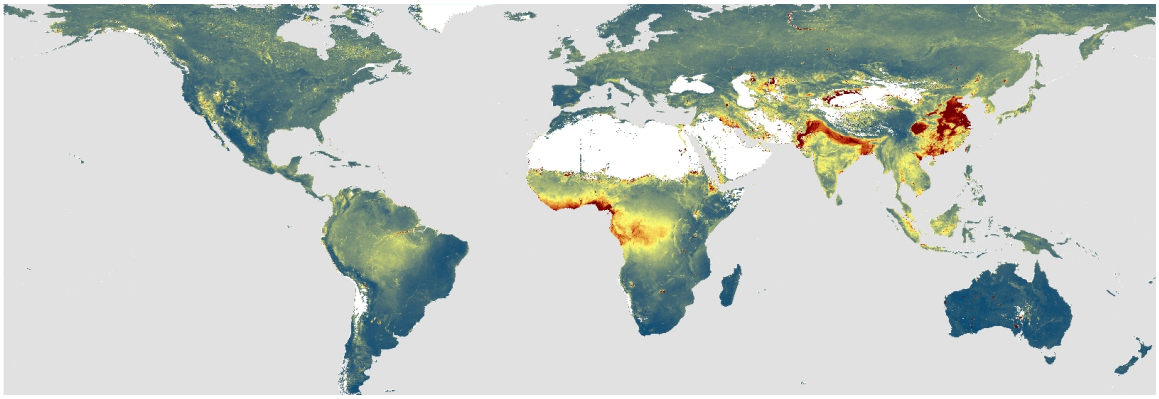
The Moderate Resolution Imaging Spectroradiometers aboard the Terra and Aqua Earth observing satellites provide daily measures of aerosol optical depth of the atmosphere at a 3km spatial resolution everywhere in the world (Levy and Hsu et al., 2015*a*,*b*). Remer, Levy, and Munchak (2013)

[1] Table A.1 reports data sources for ridership data.

Figure 2: Two maps showing AOD. Red indicates higher levels of AOD



Aerosol Optical Depth, June 1st, 2014 Terra



Aerosol Optical Depth, Average 2000-2014, Terra

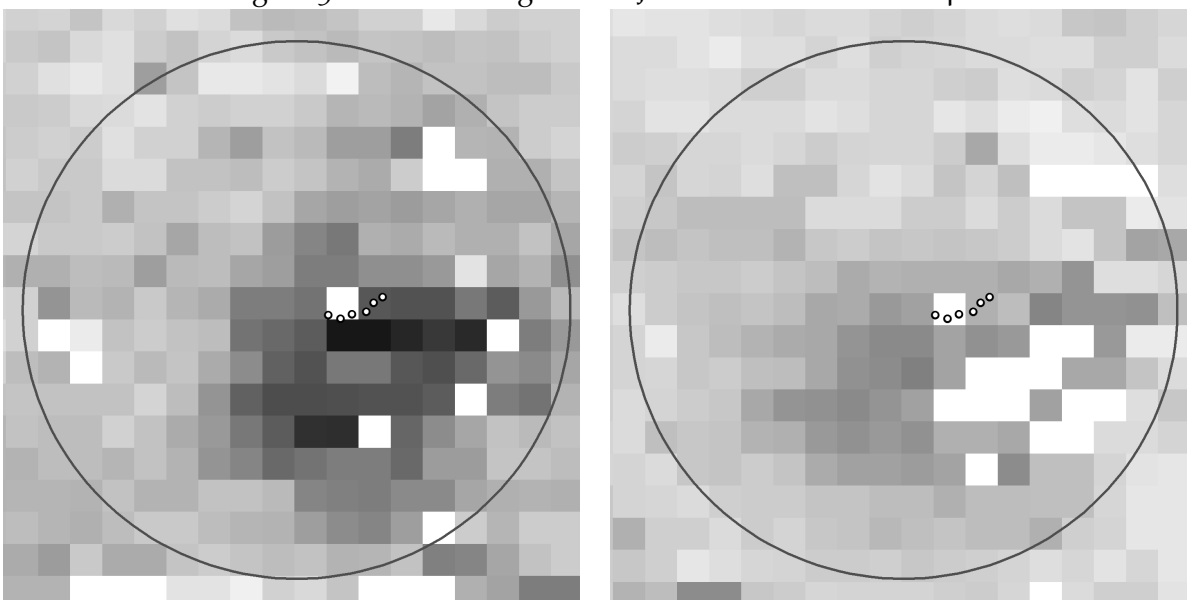
provide a description of how the AOD measure is constructed. Loosely, these instruments operate by comparing reflectance intensity in a particular band against a reference value and attributing the discrepancy to particulates in the air column.[2] The MODIS data are available for download at `ftp://ladsweb.nascom.nasa.gov/allData/6/`. Data is available in 'granules' which describe five minutes of satellite time. These granules are available more or less continuously from February 24, 2000 until December 31, 2014 for the Terra satellite, and from July 4, 2002 until December 31, 2014 for Aqua. The complete Terra and Aqua data consists of about 1.6m granules. During January of 2016, we downloaded all available granules and subsequently consolidated them into daily rasters describing global AOD. Each of these daily aggregates describes about 86m pixels covering the earth in a regular grid of 3km cells. With 28 satellite years of daily observations, this means that our monthly AOD data results from aggregating about 850b pixel-day measurements of AOD.

[2]Formally, Aerosol Optical depth is

$$\text{AOD} = -\ln\left(\frac{\text{light arriving at ground}}{\text{light arriving at top of atmosphere}}\right).$$

That is, it is a measure of the fraction of incoming light reflected by the air column before reaching the ground. Since at least zero light is reflected by the atmosphere, AOD must be positive and increasing in the share of reflected light (Jacob, 1999, p. 105). The nominal scale of AOD reported by MODIS is $0 - 5000$, although we have rescaled to $0 - 5$ for legibility, as is common in the literature.

Figure 3: AOD for Bangalore in June and December 2014



(a) Bangalore, June 2014 (b) Bangalore, December 2014

Note: Terra AOD for Bangalore in June and December of 2014. The large circle in each image has a radius of 10km and is centered on each city's central business district. Subway stations as of December 2014 are shown as black circles. Darker values indicates areas where AOD is larger and white indicates missing values.

Appendix A2 describes data processing in more detail.

Figure 2 shows the resulting images for the entire world for June 1, 2014 and for average AOD over 2000 to 2014, both from the Terra satellite. Red indicates higher AOD readings. Unsurprisingly, the figures show high AOD in India and China. Myhre *et al.* (2008) attribute high AOD over Central and Western Africa to anthropogenic biomass burning in the region. White areas indicate missing data. Because they are highly reflective, the algorithm for recovering AOD from reflectance values performs poorly over light surfaces, so missing data is common in desert regions and over snow (Levy *et al.*2013).

The MODIS instrumentation can only record AOD on cloud free days. In the June 1 image, much of the missing data reflects cloud cover, though some reflects the fact that TERRA's polar orbit brings it over most, but not all of the earth's surface each day. Because AOD reporting is sensitive to cloud cover and light surfaces, there is seasonality in the MODIS data. We see more missing data in the Northern Hemisphere in the Winter than in the Summer. The counter-cyclical Southern Hemisphere phenomena also occurs but is less dramatic.

With daily images in hand, it is straightforward to construct monthly averages. Figure 3 panel (a) illustrates the AOD data for Bangalore in June of 2014. To show scale, the large circle in this image is 10km in radius. Bangalore is noteworthy in two regards. First, there are relatively few pixels for which we have no AOD reading over the month. Second, there is a wide range in AOD

readings at this scale of observation. The corresponding picture for Delhi is entirely black. Panel (b) provides the corresponding image for Bangalore in December 2014.

We construct the monthly images presented in Figure 3 by averaging within each pixel over the course of a month. This means that, if we were to average over an area in the monthly images, a pixel which we observe only once during the month would receive the same weight as one that we observe many times. Therefore, to calculate city level monthly AOD measure, we instead average over a whole disk centered on the city for each day, and then average these city day measures, weighting by the number of pixels observed in each day. Thus, our measure of AOD within 10km of the center of a city is an average of all pixel-days of AOD readings that fall in this region during the month. We calculate this average for both satellites using disks of radius $r \in \{10\text{km}, 25\text{km}, 50\text{km}, 150\text{km}\}$ for all cities in our sample.

Table 1 provides worldwide and continental summary statistics. In 2014, the average AOD reading within 10km of a city center from the Aqua satellite was 0.42. It was higher in Asian cities, 0.56, and dramatically lower in European and North American cities. The corresponding reading from Terra is slightly higher. The top panel of Table 1 also reports AOD measurements based on disks with radius 25km centered on each city. Unsurprisingly, these larger disks have slightly lower AOD levels than the smaller and more central 10km disks. As for the 10km disks, AOD measures based on Terra are slightly higher than those for Aqua, and Asian cities are more polluted than non-Asian cities. Table A.2 reports the mean and standard deviation of AOD for each of our 43 new subway cities using the Terra satellite.

In an average month, the AOD reading for an average 10km city disk is based on 109 pixel-days for Aqua and 123 for Terra. Since the pixels are nominally 3km, if all possible pixel-days in a 10k disk were recorded over a month, we would expect about $(365/12) \times \pi \times (10km/3km)^2 = 1061$ pixels. Thus, conditional on observing one or more pixel-days, our city-month AOD values are based on measurements of about 10% of possible pixel-days. About 10% of city-months contain zero pixel-day observations and do not appear in our sample. An average AOD reading in a 25km disk is based on about 9 times as many pixel-days as the smaller 10km disks. Since the area covered by these disks is only about 6 times as large, we record a higher proportion of possible measurements in the large disks. We suspect that this reflects two factors. First, that the satellites may be worse measuring AOD over highly reflective built environments, and so do worse over more densely built up central cities. Second, pixels are included in a disk on the basis of their centroid location, and this makes it easier for the smaller disks to 'just miss' including pixels.

The second panel of Table 1 presents AOD averages for 2000 for 10 and 25km disks. Because only the Terra satellite was in operation in 2000, the second panel presents only these measures. Comparing across years, we see that variation across years is small compared to the level and compared to variation across continents. Table 1 suggests a slight downward trend in European and North American AOD, a slight upward trend in South American and Asia, and no obvious trend in the 43 cities as a whole.

Table 1: AOD in 43 new subway cities

	World	Africa	Asia	Europe	North A.	South A.
Subway openings since 2000	43	1	25	8	4	5
2014						
Av. AOD (Aqua), 10km	0.42	0.20	0.56	0.18	0.26	0.31
Av. AOD (Terra), 10km	0.45	0.23	0.59	0.21	0.28	0.3
Av. # pixels (Aqua), 10km	109.04	242.21	98.68	170.03	51.49	75.06
Av. # pixels (Terra), 10km	123.64	255.94	114.41	183.55	56.84	96.26
Av. AOD (Aqua), 25km	0.37	0.17	0.51	0.15	0.19	0.23
Av. AOD (Terra), 25km	0.41	0.20	0.55	0.18	0.20	0.26
Av. # pixels (Aqua), 25km	989.77	2269.21	914.96	1436.29	672.48	637.03
Av. # pixels (Terra), 25km	1080.92	2418.59	1016.73	1522.35	568.57	820.31
2000						
Av. AOD (Terra), 10km	0.43	0.27	0.54	0.28	0.34	0.29
Av. AOD (Terra), 25km	0.38	0.23	0.49	0.24	0.22	0.23

Figure A.1 illustrates the extent of satellite coverage for our sample of 43 subway cities. Panel (a) shows the count of cities by month for which we observe AOD in a 10km disk surrounding the city's center for each of the two MODIS satellites. Panel (b) plots mean AOD for 10km disks centered on the central business districts of our 43 subway cities. Both figures show a strong seasonal pattern in AOD. This reflects seasonal variation in cloud cover and motivates our use of city-by-calendar month indicators in all of our main regressions.

Together the two panels of Figure A.1 suggest that a relationship exists between the extent of coverage and the level of AOD. In fact, a regression of average AOD in a city-month disk on the count of pixel-days used to calculate that average reveals a slightly positive relationship. We conjecture that this reflects the fact that the air is cleaner in rainy places where cloud cover is more common. Regardless of the reason, we have experimented broadly with sampling rules that reduce the importance of city-months for which AOD data is sparse and with controlling for the count of pixel-days used to construct each city-month average. In most of our regression results we control for the number of pixel-days used to construct each city-month AOD observation.

C *Other control variables*

In Section 4 we validate the relationship between the MODIS data and ground based measurements in our sample. Consistent with the large related literature, which we describe in Section 4, we find that local weather conditions are important determinants of MODIS AOD. Given this, we construct several controls for city-month weather conditions. The CRU gridded dataset from Harris *et al.* (2014) provides high-resolution monthly climatic data describing cloud cover percentage, frost day frequency, mean temperature, precipitation, and vapour pressure. We use these data to

calculate monthly and annual averages of these variables over disks centered on each city. These are our 'climate controls'.

We also include city population and country GDP per capita to characterize the level of economic activity of each city. Our population data comes from the United Nation's 2014 Revision of the World Urbanization Prospects (DESA Population Division, 2014). These data describe annual population counts for all urban agglomerations with populations exceeding 300,000 at any time between 1950 and 2014 and also provide coordinates for the centers of all of the cities they describe. With a few exceptions that we adjusted by hand on the basis of lights at night data, we use these coordinates for the centers of all of our cities. We use the Penn World Tables to obtain annual measures of country GDP (Feenstra, Inklaar, and Timmer, 2015) for all cities in our sample.[3] We interpolate these annual GDP and population data to construct monthly measures.

Table A.2 reports population and country level GDP per capita for each city in the 12th month after the subway system opening. Cities that open subways are large, their average population is 3.7m, and tend to be in middle or high income countries.

4. Aerosol Optical Depth versus ground based measurements

A series of papers compare measures of AOD to measures of particulate concentration from surface instruments (e.g. Gupta, Christopher, Wang, Gehrig, Lee, and Kumar, 2006, Kumar, Chu, and Foster, 2007, Kumar, Chu, Foster, Peters, and Willis, 2011). In particular, Kumar *et al.* (2007) examines the ability of AOD to predict particulates in a set of large cities, several of which are subway cities.

Broadly, this literature concludes that AOD is a good measure of airborne particulates, with two caveats. First, satellite reports of AOD describe daytime average conditions over a wide area at the particular time the satellite passes overhead, while ground based instruments record conditions at a particular location, often over a period of hours. This causes an obvious divergence between satellite and ground based measures.

In addition, ground based instruments report the concentration of dry particulates, while the satellite based measure has trouble distinguishing water vapor from other particles. This suggests that some method of accounting for differences in climate will be important when we examine the relationship between subways and AOD.

As a direct check on our AOD data, we use World Health Organization data (WHO, 2016a) describing average annual PM10 and PM2.5 concentrations ($\mu g/m^3$) in cities where ground-based pollution readings were available. We successfully match 143 such cities to our subway cities data. Of the 143, 20 report readings during four years, 80 during three years, 28 during two years and 15 in only one year. The readings span the 2003-2014 period, and not all city-years record

[3] Two of our subway cities, Algiers and Dubai, lie in countries for which the Penn World Tables do not provide GDP (Algeria and UAE). For these two cities, we use country level GDP information from the World Bank.

Table 2: The relationship between AOD and ground-based particulate measures

	PM10			PM2.5		
	(1)	(2)	(3)	(4)	(5)	(6)
Terra AOD	122.85***	99.62***	75.26***	61.58***	55.12***	42.53***
	(10.09)	(9.33)	(9.24)	(8.18)	(6.11)	(6.10)
Constant	-1.26	59.37	123.66***	-1.11	24.74	56.59**
	(2.99)	(40.68)	(38.28)	(2.25)	(25.18)	(23.43)
Mean dep. var.	46.17	46.17	46.54	19.62	19.62	19.66
Mean indep. var.	0.39	0.39	0.39	0.34	0.34	0.34
R-squared	0.54	0.75	0.81	0.51	0.73	0.81
Cities	140	140	138	85	85	84
N	316	316	311	227	227	225
Aqua AOD	124.49***	101.72***	76.15***	63.47***	56.74***	43.11***
	(10.52)	(10.13)	(10.08)	(8.39)	(6.05)	(6.43)
Constant	2.06	51.80	119.35***	0.10	16.42	50.45**
	(2.84)	(43.41)	(39.95)	(2.10)	(27.28)	(25.43)
Mean dep. var.	46.17	46.17	46.54	19.62	19.62	19.66
Mean indep. var.	0.35	0.35	0.36	0.31	0.31	0.31
R-squared	0.52	0.74	0.80	0.48	0.71	0.80
Cities	140	140	138	85	85	84
N	316	316	311	227	227	225

*Note: Aerosol Optical Depth is the mean value in a 10km radius disk around the city center. Columns (1) and (4) have no control variables. Columns (2) and (5) add climate controls and continent dummies. Columns (3) and (6) add controls for city population and country GDP. Robust standard errors in parentheses. Stars denote significance levels: * 0.10 ** 0.05 *** 0.01.*

both PM10 and PM2.5 . Averaging monthly AOD values to calculate yearly averages, we obtain 316 comparable city-years for PM10 and 227 comparable city-years for PM2.5 .

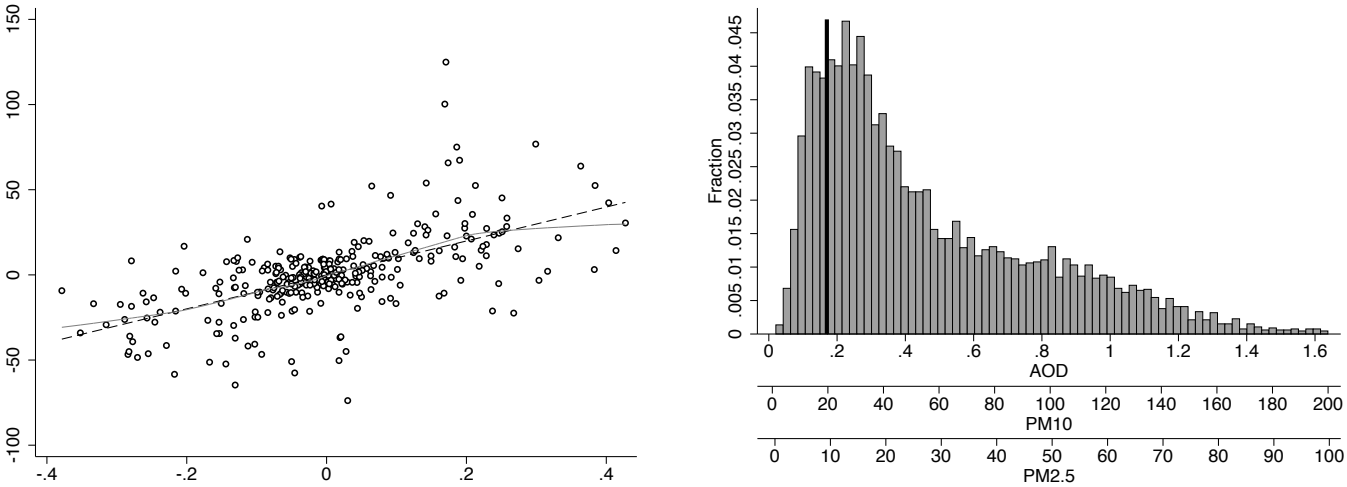
To compare the WHO ground based annual measures of particulates to annual averages of MODIS AOD measurements in subway cities, we estimate the following regressions

$$\mathrm{PM}_{yit} = \alpha_0 + \alpha_1 \mathrm{AOD}_{it} + \mathrm{controls}_{it} + \epsilon_{it},$$

where $y \in \{2.5,10\}$ is particulate size, i refers to cities and t to years for which we can match WHO data to our AOD sample.

Table 2 reports results. The upper first column presents the results of a regression of the WHO measure of PM10 on annual average Terra AOD within 10km of a subway city center. There is a strong positive relationship between the two quantities and the R^2 of the regression is 0.54. The AOD coefficient of 122.85 in column 1 means that a one unit increase in AOD maps to a 122.85 $\mu g/m^3$ increase in PM10 . From Table 1, we see that Terra 10k readings for North America decreased by 0.06 in subway cities between 2000 and 2014. Multiplying by 122.85 gives a $7\mu g^3$ decrease. By

Figure 4: AOD versus PM



(a) (b)

Note: (a) Plot showing residualized PM10 and AOD, together with linear trend and locally weighted regression line. (b) Histogram of city-months by AOD, PM10 *and* PM2.5 *.* PM10 *and* PM2.5 *axes rescaled from AOD using columns 1 and 4 of Table 2. Black vertical line indicates WHO threshold level for annual average* PM10 *exposure (WHO, 2006).*

contrast, according to US EPA historical data, during this same period US average PM10 declined from 65.6 to 55.0 μg/m^3, or about a 10 unit decrease.[4] Since Table 1 reports AOD for just the four cities in North America with new subways, while the EPA reports area weighted measures for the US, this seems reasonably close.

All specifications reported in Table 2 assume a linear relationship between AOD and PM. Figure 4(a) plots residuals of regressions of PM10 and AOD on all controls used in column 2 of Table 2, along with linear and locally weighted regression lines. This graph illustrates both how closely the two variables track each other and how close to linear is the relationship between them.

In column 2, we conduct the same regression but include linear and quadratic terms in our annual climate variables, counts of AOD pixel-days, and continent indicators. The coefficient on AOD drops from 122.85 to 99.62, and the R^2 increases to 0.75. In column 3, we add controls for city population and country GDP. This decreases the coefficient on AOD to 75.26 and increases the R^2 of the regression to 0.81. Columns (4) to (6) replicate (1) to (3) but use PM2.5 as the dependent variable. These regressions have greater predictive power, with qualitatively similar results. The AOD coefficients are about half as large, which is consistent with the PM10 to PM2.5 conversion factors used by the World Health Organisation (WHO, 2016a). Finally, the lower panel of Table 2 shows analogous regressions for Aqua. We conduct, but do not present, corresponding regressions where our AOD measure is based on a 25km ring around the city center instead of 10km. All results

[4] `https://www.epa.gov/air-trends/particulate-matter-pm10-trends`, accessed April 3, 2017.

are broadly similar.[5]

Recall that the ground-based instruments and MODIS, in fact, measure something different. Ground-based instruments measure pollution at a point over an extended period of time. Remote sensing measures particulates across a wide area at an instant. Given this difference, the extent to which the two measures appear to agree seems remarkable.

In addition to validating the use of remotely sensed AOD, Table 2 provides a basis for translating our estimates of the relationship between subways and AOD into a relationship between subways and PM2.5, or PM10. To illustrate this process, and to help to describe our data, Figure 4(b) provides a histogram of the 12,169 city-months used for our main econometric analysis. The figure provides three different scales for the horizontal axis. The top scale is the raw AOD measure. The second two axes are affine transformations of the AOD scale into PM10 and PM2.5 based on columns 1 and 4 of Table 2. For reference, the black line in the figure gives the World Health Organization recommended maximum annual average PM10 exposure level (20 μg/m^3).

5. The relationship between subway system openings and AOD

In our primary econometric exercise, we examine changes in AOD around the time that a city opens its subway system. Let $i = 1,...,I$ index subway cities and t index months between February 2000 and December 2014. Thus, AOD_{it} denotes AOD in city i at time t. We are interested in changes to AOD in the months around a system opening. In fact, we usually observe each city twice in each month, once with each of the two satellites, but suppress the satellite subscript for legibility.

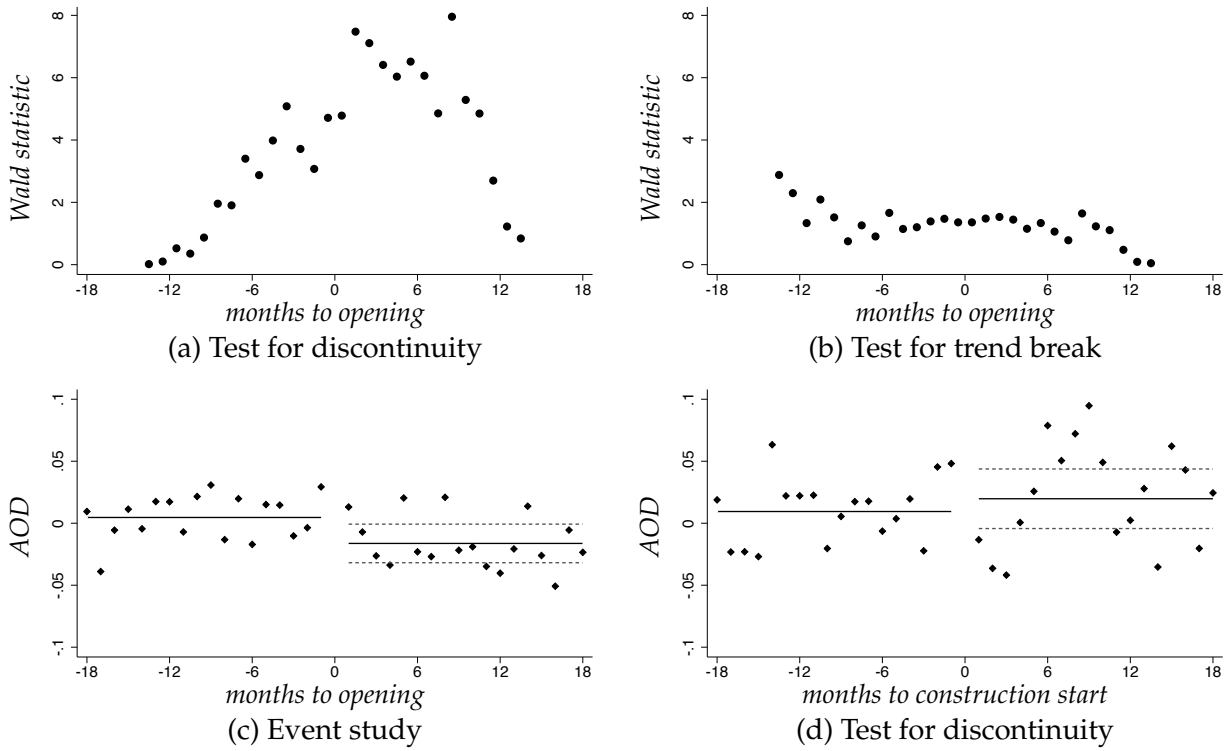
If city i opens its subway in month t', then define $\tau_{it} = t - t'$. That is, τ_{it} is 'months since the subway opened', with months before the opening taking negative values. Let k describe the window over which we analyze AOD, i.e., $\tau_{it} \in \{-k,...,0,...,k\}$. We will most often be interested in the case of $k = 18$, that is, the 37 month period extending from 18 months before until 18 months after a subway opening. There are 39 cities that open a subway for which this entire window falls

[5] We note that the results in Table 2 are quite different from those on which the 2013 Global burden of disease estimates are based (Brauer *et al.* 2015). In particular, they estimate

$$\ln(\text{PM2.5}) \approx 0.8 + 0.7 \ln(\text{AOD}).$$

Comparing with Table 2, we see that these coefficient estimates are quite different. The difference reflects primarily our use of the level of PM2.5, rather than its logarithm, as the dependent variable. We also use the level of AOD rather than its logarithm as the explanatory variable. Since AOD is typically around 0.5, this turns out not to be important. Finally, our sample describes a different and more urban sample of locations, relies on annual rather than daily data, and measures AOD using just MODIS data rather than an average of MODIS and a measure imputed using a climate model and ground based emissions release information. We prefer the formulation in Table 2 to that in Brauer *et al.* (2015) for three reasons. First, AOD is already a logarithm (see footnote 2), so the Brauer *et al.* specification uses the logarithm of a logarithm as its main explanatory variable. Second, mortality and morbidity estimates are typically based on levels of pollutants, not on percentage changes, so the dependent variable in our regressions is more immediately useful for evaluating the health implications of changes in AOD. Finally, we control for weather conditions, which appears to be important. In any case, the R^2 in both studies is of similar magnitude.

Figure 5: Break-tests and event studies the 18 months before and after subway openings and the start of construction



Notes: (a) Plot of Wald statistics for tests of a regression intercept discontinuity at time τ. Test statistics calculated in regressions that also control for a satellite indicator, year-by-continent indicator variables, city-by-calendar month indicators, AOD pixel-days and linear and quadratic terms of monthly climate variables, and linear and quadratic terms in country level GDP and city level population. (b) Plot of Wald statistics for tests of a trend break at time τ conditional on a discontinuity in the mean level of AOD at $\tau = 1$. Other details are the same as in Panel (a). (c) Event study during 18 months before and after subway openings. (d) Event study during 18 months before and after start of subway construction.

within the February 2000 to December 2014 range of our AOD data. In order to keep a constant sample of cities for each month, we most often consider this set.

Now define the following families of indicator variables,

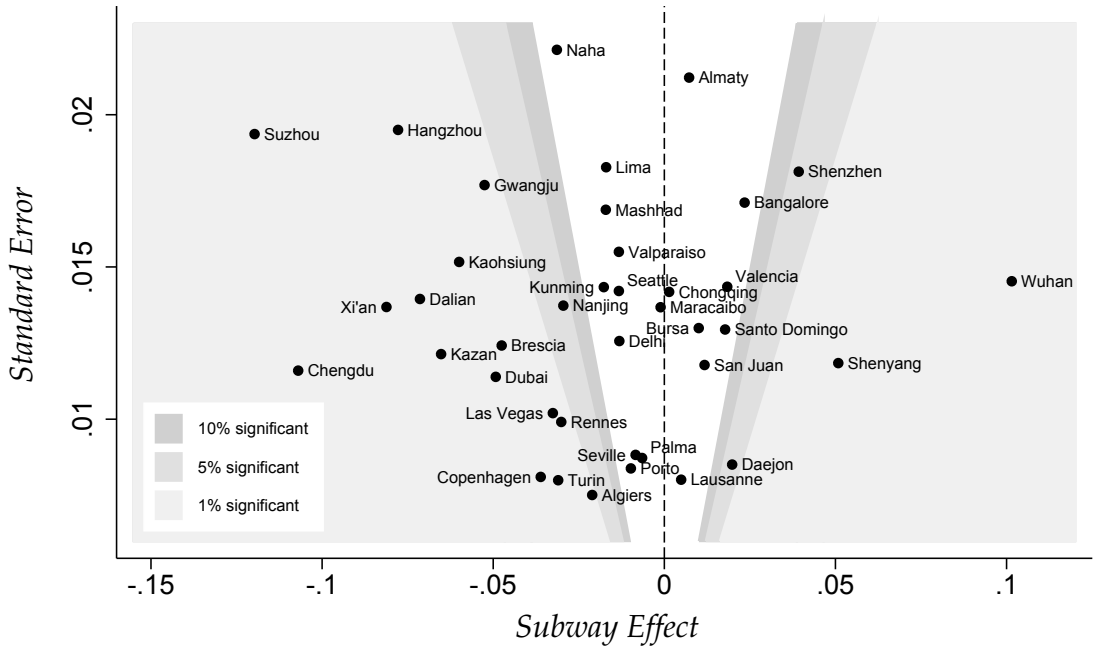
$$D_{it}(j) = \begin{cases} 1 & \tau_{it} = j \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

$$D_{it}(j,j') = \begin{cases} 1 & \tau_{it} \in \{j,\ldots,j'\} \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

$$D_{it}(\tau \lesseqgtr j') = \begin{cases} 1 & \tau_{it} \lesseqgtr j' \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Equation (1) describes indicators for sets of city-months that are the same number of months away from the month when their subway system opens. Equation (2) describes indicators for a set of step functions beginning j and ending in j' months from the subway opening month. Finally, equation

Figure 6: Heterogeneity of the effect of subway opening on AOD



Notes: Illustration of 39 city specific subway effects. x axis is the estimated treatment effect, y is the standard error of the estimated treatment effect. Region in white contains estimates that are not significantly different from zero.

(3) describes indicators for city-months that are far enough from the subway opening date that they fall outside our study window.

Sections 3 and 4 indicate the following patterns in the data. The AOD data are seasonal and the pattern of seasonality varies across the globe. The AOD measurements reported by Terra are systematically larger than Aqua, but otherwise the two satellites track each other very closely. Remotely sensed AOD may confound water vapor with anthropogenic particles. City-month AOD averages are slightly increasing in the number of pixel-days on which they are based. Finally, there are long run trends in pollution, and these trends are probably different on different continents.

Given this, we include the following controls in our regressions; a satellite indicator, year-by-continent indicators, and city-by-calendar month indicators. We also include the count of AOD pixel-days used to calculate the city-month AOD and linear and quadratic terms of our monthly climate variables. Finally, the control set sometimes includes linear and quadratic terms in country level GDP and city level population. Subject to the change from annual to monthly observations, these control variables closely correspond to those we used in our investigation of the relationship between PM10 , PM2.5 and AOD in Table 2.

Following Andrews (1993), Hansen (2000), and Andrews (2003), we check for a structural break in the value of AOD around the time of system opening by estimating a series of regressions of AOD on a step function as we allow the timing of the step to traverse the study period. The

Table 3: Subway opening and AOD for the 18 month period post system opening

	(1)	(2)	(3)	(4)	(5)	(6)
1-18 months post	-0.0205**	-0.0209**	-0.0201**	-0.0210**	-0.0213**	-0.0223**
	(0.00931)	(0.00946)	(0.00883)	(0.00895)	(0.00815)	(0.00865)
	[0.0266]	[0.0266]	[0.0200]	[0.0133]	[0.0133]	[0.0200]
city pop.& GDP	No	Yes	No	Yes	No	Yes
city-level trends	No	No	Yes	Yes	No	No
city-level pre/post trends	No	No	No	No	Yes	Yes
Mean dep. var.	0.457	0.457	0.457	0.457	0.457	0.457
R-squared	0.833	0.834	0.838	0.838	0.839	0.839
Number of events	39	39	39	39	39	39
N	12,169	12,169	12,169	12,169	12,169	12,169

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. p-values on the coefficient of interest using the wild cluster bootstrap procedure in Cameron, Gelbach, and Miller (2008) are reported in square brackets. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month×city indicators, year×continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

relevant regressions are,

$$\begin{aligned} AOD_{it} = \alpha_0 + \alpha_{1j}\, D_{it}(j,k) \\ + \alpha_2\, D_{it}(\tau < -k) + \alpha_3\, D_{it}(\tau > k) + \text{controls}_{it} + \epsilon_{it} \\ \text{for all } j \in \{-0.75k,...,0,...,0.75k\} \end{aligned} \tag{4}$$

For our main analysis we set $k = 18$. This window length strikes a balance between maintaining the set of cities from which we identify our coefficient of interest and having a long analysis window. Thus, we use a 37 month study window and estimate equation (4) for each month in $j \in \{-14,...,14\}$ with errors clustered at the city level. We then calculate a Wald test of $\alpha_{1j} = 0$ for each j. By including pre- and post-period indicators in these regressions we use all city-months in our sample to estimate city-by-calendar month indicators, continent-by-year indicators and climate variables, while only using AOD variation near the subway opening date to identify the effect of subways on AOD. Panel (a) of Figure 5 plots these test statistics.

The figure shows a clear pattern. Wald statistics increase from a low level to a peak right after the subway opening, before quickly falling. That is, this plot suggests that the subway system opening leads to a break in the AOD sequence, and that this break occurs around the time when

the system opens.[6]

On the basis of Panel (a) of Figure 5, we assume a break in AOD levels that coincides with the first full month of subway system operations, i.e., $\tau = 1$. Conditional on such a break, we next check for a change in the trend of AOD associated with subway openings. We proceed much as in our test for a break, but instead look for a change in trend around the time of a subway opening. Formally, this means estimating the following set of regressions,

$$\begin{aligned} AOD_{it} = \alpha_0 + \alpha_1 \tau_{it} + \alpha_{2j} \tau_{it} D_{it}(j,k) + \alpha_3 D_{it}(1,k) \\ + \alpha_4 D_{it}(\tau < -k) + \alpha_5 D_{it}(\tau > k) + \text{controls}_{it} + \epsilon_{it} \\ \text{for all } j \in \{-0.75k,...,0,...,0.75k\} \end{aligned} \tag{5}$$

As before, we estimate the regression (5) for each month in $j \in \{-14,...,14\}$ with errors clustered at the city level and calculate the Wald test for $\alpha_2 = 0$ for each regression.[7] Panel (b) of Figure 5 plots these Wald statistic values as j varies.[8] Thus, conditional on a step at $\tau = 1$, subways openings do not seem to cause a change in the trend of AOD in a city.

Figure 5 suggests that subway openings cause a one time shift in a city's level of AOD, and no other change in the evolution of AOD. Given this evidence for the existence and timing of an effect of subway openings on the evolution of AOD, we turn to estimating its size.[9]

To illustrate patterns in the data, we also estimate,

$$\begin{aligned} AOD_{it} = \alpha_0 + \sum_{j \in \{-k,...,k\} \setminus \{0\}} \alpha_{1j} D_{it}(j) \\ + \alpha_2 \, D_{it}(\tau < -k) + \alpha_3 \, D_{it}(\tau > k) + \text{controls}_{it} + \epsilon_{it} \end{aligned} \tag{6}$$

Control variables are the same as those used in our test for a structural break. Since the indicator for $\tau_{it} = 0$ is omitted, the α_{1j} measure mean difference in AOD for city-months with $\tau_{it} = j$ from those with $\tau_{it} = 0$. Figure 5 plots as dots the 36 monthly values of α_{1j} that result from conducting regression (6) with $k = 18$. The vertical axis of this figure indicates AOD. From Table 1, sample mean AOD is about 0.44, so monthly variation in AOD is small relative to the mean level. Inspection of Figure 5 suggests a drop in AOD around the time of a system opening.

We are concerned that AOD responds not to subway openings, but to subway construction. This might occur for at least three reasons. First, construction is dusty, and so ending it might

[6] Andrews (2003) gives (asymptotic) critical values for the test statistic values we have just generated, a 'sup-Wald' test for $\alpha_{1j} = 0$ for all j. For our case, where the break in question affects only one parameter and where we trim 25% from the boundaries of the sample, the 5% critical value of this statistic is 7.87, less than the largest value that we observe for the Wald statistic in the months after a system opening. With this said, our estimation framework differs from the one for which this test statistic is derived in several small ways, and so we regard this test with some caution.

[7] An alternative would be to simultaneously search for locations of the break and trend break. Hansen (2000) argues that sequential searching, as we do, arrives at the same result.

[8] All values are well below the 10% critical value of 6.35 given in Andrews (2003). Again, our framework differs from the framework under which this test statistic is derived so this test should be regarded with caution.

[9] Hansen (2000) provides an elegant method for estimating the location and size of the break along with confidence intervals. This method assumes a balanced panel, and so we have not been able to apply it to our data.

reduce pollution independent of the effect of the opening of the subway. Second, construction is expensive and labor intensive. AOD might decrease in response to an overall decrease in economic activity when construction ends. Finally, subways might occupy space otherwise used for cars and thereby affect traffic through a channel unrelated to the subway's intended function.

To assess the importance and plausibility of these hypotheses, we collect data describing the start date of construction. These data, detailed in Table A.2, indicate that on average construction began almost seven years prior to opening, with considerable variation around this mean. These data allow us to replicate the event study described in equation 6 around the beginning of construction. The bottom right panel of Figure 5 presents the results and does not support the idea that the start of subway construction affects city AOD levels. While it is easy to come up with reasons about why some event related to subway construction rather subway openings might affect AOD, the data do not offer immediate support for such hypotheses.

To estimate the size of the post-subway drop in AOD, and to investigate its robustness, we conduct regressions of the following form,

$$\begin{aligned} AOD_{it} = \alpha_0 + \alpha_1 D_{it}(1,k) \\ + \alpha_2 D_{it}(0) + \alpha_3 D_{it}(\tau < -k) + \alpha_4 D_{it}(\tau > k) + \text{controls}_{it} + \epsilon_{it}. \end{aligned} \tag{7}$$

This is a regression of AOD on an indicator for post-opening and controls. As before, we include pre- and post-period indicators in order to isolate AOD variation for city-months close to the time of their subway opening. We include a dummy variable for city-months when $\tau_{it} = 0$. This month is 'partly treated' and this specification ignores AOD variation for these partly-treated city-months. The intercept, α_0, gives the conditional mean AOD over the city-months where $\tau_{it} \in \{-18,..., -1\}$. The coefficient of interest is α_1, the difference in conditional mean AOD between city-months with $\tau_{it} \in \{-18,..., -1\}$ and those with $\tau_{it} \in \{1,...,18\}$.

Table 3 presents results. In column 1 of Table 3 we conduct the regression given in equation (7) with our basic set of controls.[10] We see that the effect of subways on AOD in the 18 months following a subway opening, α_1, is about -0.02 and is different from zero at the 5% level. Column 2 adds country level GDP and city population. Column 3 adds city specific trends. Column 4 adds country level GDP and population and city specific trends. Column 5 adds city specific trends and allows for a trend break in each city at time zero. Column 6 adds country level GDP and population to the specification of column 5. The estimated effect on AOD of subway opening is stable across these various specifications. The difference between the smallest and largest coefficients (columns 3 and 6) is 0.002. This is about 25% of the standard error of the most precise estimate in column 5. Panel (c) of Figure 5 shows the magnitude of the AOD coefficient from column 2, as well as 95% confidence bounds.

[10]For reference, we observe 39 cities where subways open, for 12 month per year, for 14 years with two satellites. Multiplying gives 13,104. This slightly exceeds the size of our main regression sample because the Aqua satellite begins operation in 2002 rather than 2000 for Terra, and because of a few city-months for which a satellite reports zero AOD pixels.

All regressions in Table 3 include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month×city indicators, year×continent indicators, pre- and post-period indicators. Removing climate controls or AOD pixel count does not have an important effect on the subway coefficient or the regression R^2. We also experimented broadly with sampling rules that reduce the importance of city-months for which AOD data is sparse and find that it does not qualitatively affect our results. However, calendar month×city indicators are important. If we omit these indicators, all else equal, the R^2 of our regressions drops from about 0.83 to about 0.49, while the coefficient for subway openings about doubles in the first four columns of Table 3, but is not distinguishable from zero when we include city specific trend with a break at zero in columns 5 and 6. That the correlation between subway openings and AOD does not depend on weather seems unsurprising. Since the number of AOD pixels primarily reflects cloud cover, the same intuition extends to the pixel count control. The importance of calendar month control is also expected. City level seasonality is an important determinant of AOD levels.

Cameron, Gelbach, and Miller (2008) find that asymptotic standard errors may not approximate exact finite sample standard errors in samples containing 30 or fewer clusters. With 39 cities/clusters our main sample is not much larger. To confirm that the asymptotic standard errors reported in Table 3 approximate exact values, we implement the wild-cluster bootstrap procedure recommended in Cameron *et al.* (2008). We report wild cluster bootstrap p-values using 300 bootstrap replications in square brackets in Table 3. The statistical significance of the estimates is unchanged when we use the wild cluster bootstrap instead of the asymptotic standard errors. Given this, we report only asymptotic errors for the rest of our results.

Table 3 strongly suggests that opening a subway network decreases a city's AOD by about 0.02 over the 18 month period following the opening. The mean value of AOD across cities in our sample before subways open is 0.49, so this 0.02 subway effect on AOD represents a 4% reduction on average.

Regression results so far focus on average effects over the 18 months before and after a subway opening. Table A.3 refines the results of Table 3 by decomposing the study window into 6 month bins. These estimates are broadly consistent with inspection of panel (c) of Figure 5 and Table 3. Subways have a negative effect on AOD during each six month period following a subway opening. These effects are largest during the period from 7-12 months after an opening. Unsurprisingly, our estimates are somewhat imprecise, and only the effect during the 7-12 months post-opening is distinguishable from zero at conventional confidence levels. Estimates of the subway effect on AOD in the two pre-periods are dramatically smaller than post-periods, but are estimated with about the same precision. In particular, and reassuringly, they are not distinguishable from zero. In columns 5 and 6 of Table 3 we see that our estimates are robust to the inclusion of city specific trends, before and after the opening. Together with the results of Table A.3 this suggests that our results are not driven by confounding trends in AOD around the time of subway openings.

Table 3 presents estimates of the average effect of subways on a city. It is natural to suspect that, in fact, subways do not affect all cities in the same way. To investigate this possibility, we estimate a version of equation 7 where we allow the effect of subways to vary across cities. Figure 6 summarizes these results. The horizontal axis gives the magnitude of the subway effect for each city, and the vertical axis the corresponding standard error. Shaded regions indicate whether the city effect in question is distinguishable from zero at usual confidence levels. This figure suggests considerable heterogeneity in how cities respond to subways.

We conduct a number of exercises in order to discern a pattern in this heterogeneity. Table A.4 investigates whether subways have different effects on different types of cities. In this table, we replicate the results of column 2 of Table 3, but add an interaction between the post-subway indicator and a particular city characteristic, along with the characteristic in question. In order, the interactions considered are; an indicator for cities in the top half of the city size distribution in 2000, an indicator for cities in the bottom half of the country income distribution in 1990, an indicator for above median average rainfall, an indicator for cities whose subway opening included more than the median number of stations, indicator for cities in Asia, and finally, an indicator for cities in the top half of the AOD distribution in the year prior to their subway opening.

The interaction term has only a small effect on the main treatment effect and the interaction term is small and indistinguishable from zero when we consider heterogeneity of effects by city size, country income level, mean precipitation, and initial subway size (Columns 1-4 in the table). Heterogenous effects are more evident when we differentiate between Asian cities and the rest (column 5), and indicate that Asian cities experience larger pollution reductions with the introduction of their subway systems (-0.025 vs -0.015), although the interaction term is not statistically different from zero. Column 6 distinguishes between high and low pollution cities. As would be expected, the estimates suggest that pollution effects are larger (-0.029 vs -0.012) among initially more polluted cities. In fact, this specification indicates that pollution reduction effects can only be differentiated from zero for initially higher pollution cities.

A *Longer time horizons*

Subways are durable and their effects probably extend over decades. Hence, it is of interest to extend our estimates of the effects of subways to the longest possible horizon that our data permit. Unfortunately, considering a longer treatment period requires that we degrade our research design in one of two ways. As we consider longer treatment periods we must either allow later post-treatment effects to reflect a decreasing set of cities, opening the door to confounding composition with subway effects, or else restrict attention to progressively smaller samples of cities, reducing precision and raising questions of external validity. In spite of this, the importance of obtaining estimates over a time horizon that more nearly approximates the planning horizon suggests that such estimates will be useful, even though we have less confidence in them.

In Table 4 we continue to consider a pre-treatment period beginning 18 months before an opening, but consider longer post-treatment periods. In columns 1 and 2 we consider two years after an opening, and allow the subway effect to vary by year using a specification that is otherwise the same as we used in Table 3. As in Table 3, the two columns differ only in that column 2 includes controls for city population and country level GDP, while column 1 does not. We see that the one year effect is about -0.02, statistically identical to our estimate of the 18 month effect in Table 3. Point estimates of the second year effect are slightly smaller and are estimated with about the same precision as the 1 year effect. We can reject neither the hypothesis that the second year effect is zero nor that it is the same as the one year effect. Panel (b) of Table 4 estimates the average subway effect over the two year post-period considered in the first two columns. Unsurprisingly, this average effect is statistically different from zero, but not from -0.02. In order to extend our sample to two years post-treatment with a constant sample of cities in all months, we lose a city and the first two columns of Table 4 reflect 38 subway openings instead of 39.

In columns 3 and 4 of Table 4 we extend the post-treatment period to 36 months. Each of the three post treatment years are negative and estimated precisely enough that they can be distinguished from zero. Point estimates are slightly increasing in magnitude with time from the system opening, although the change is not large relative to the magnitude or precision of the estimates. To consider the three year post-treatment period with a constant sample of cities in each month, we restrict our sample to 35 cities. Thus, the slight change in the year 1 and 2 treatments from columns 1 and 2 to columns 3 and 4 may reflect the change in sample. Finally, columns 5 and 6 consider a four year post-treatment period. These columns indicate slightly larger effects than in the other columns, and suggest that the effect of subways increases over time, although the precision of these estimates does not allow us to reject the hypothesis that the effect is constant in all post-treatment years. To preserve a constant sample of cities over this longer horizon, we restrict our sample to only 28 subway openings. Again, the average effect presented in panel (b) is not distinguishable from -0.02.

Table 5 extends the analysis horizon by dropping the requirement that the sample of cities is constant throughout the window of analysis and allows us to show estimates of the effect of subways over the five to eight years after a system opening. To begin, columns 1 and 2 of Table 5 look at the effect of subways for a five year post-treatment period following the specification of column 2 of Table 3, except for the longer treatment period. For reference, column 1 restricts attention to set of 26 cities for which we observe the entire post-treatment period.[11] In column 1 we see that all yearly effects are negative, and, except for the very large effect in the fifth year, not far from our baseline estimate of about -0.02. Switching to the larger sample of 39 cities in column 2, we see that coefficient magnitudes and standard error decrease slightly. In columns 3 to 5, we consider progressively longer treatment periods. Panel (b) reports the average effect over

[11] In Table A.2 these 26 cities range from Rennes to Dubai.

Table 4: Longer term effects

	(1)	(2)	(3)	(4)	(5)	(6)
Panel a.						
1-12 months post	-0.0203*	-0.0208*	-0.0186*	-0.0185*	-0.0222*	-0.0209*
	(0.0102)	(0.0103)	(0.0108)	(0.0108)	(0.0116)	(0.0109)
13-24 months post	-0.0127	-0.0129	-0.0178**	-0.0174*	-0.0250***	-0.0224**
	(0.0097)	(0.0010)	(0.0087)	(0.0091)	(0.0090)	(0.00865)
25-36 months post			-0.0271**	-0.0260**	-0.0366***	-0.0324***
			(0.0103)	(0.0098)	(0.0107)	(0.0095)
37-48 months post					-0.0316*	-0.0255*
					(0.0158)	(0.0129)
City pop. and GDP	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.462	0.462	0.443	0.443	0.416	0.416
R-squared	0.833	0.834	0.831	0.832	0.829	0.830
Number of events	38	38	35	35	28	28
N	11,841	11,841	10,881	10,881	8,863	8,863
Panel b.						
average post-period	-0.0167*	-0.0171*	-0.0209**	-0.0204**	-0.0273***	-0.0243***
	(0.0087)	(0.0089)	(0.0081)	(0.0079)	(0.0089)	(0.0072)

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

the whole relevant post-treatment period. In every case this average effect is quite close to our baseline specification in column 1 of table 3.

Inspection of these results reveals three patterns. First, across specifications, point estimates are within 1.64 standard errors of -0.02 in almost all cases. Second, only the first, third and fifth year effect are significantly different from zero and the fifth year effect seems anomalously large relative to the other years. Finally, there is no obvious trend in coefficient magnitudes with time from opening, while in all columns except 1, standard errors increase monotonically, or very nearly so, with time from opening. That is, unsurprisingly, as the subway opening is more remote, its effects are progressively more difficult to detect, and beyond 5-6 years we are unable to distinguish annual subway effects from zero. The lower panel in the table reports mean effects across post-treatment years and points to surprisingly stable estimates even as the number of years post treatment are extended (-0.019 in column 2 versus -0.018 in column 5). These results are consistent with a persistent AOD decrease of about 0.02 that is gradually overwhelmed by noise or with a transitory

Table 5: Even longer horizon

	(1)	(2)	(3)	(4)	(5)
Panel a.					
1-12 months post	-0.0172	-0.0200*	-0.0199*	-0.0194*	-0.0195*
	(0.0109)	(0.0100)	(0.0100)	(0.0099)	(0.0097)
13-24 months post	-0.0218*	-0.0120	-0.0119	-0.0108	-0.0110
	(0.0107)	(0.0097)	(0.0010)	(0.0099)	(0.0099)
25-36 months post	-0.0290***	-0.0192*	-0.0192*	-0.0177*	-0.0180
	(0.0101)	(0.00997)	(0.0104)	(0.0104)	(0.0109)
37-48 months post	-0.0243	-0.0177	-0.0176	-0.0158	-0.0162
	(0.0155)	(0.0137)	(0.0142)	(0.0142)	(0.0146)
49-60 months post	-0.0434**	-0.0369***	-0.0368**	-0.0346**	-0.0351**
	(0.0165)	(0.0132)	(0.0137)	(0.0138)	(0.0144)
61-72 months post			-0.0237*	-0.0215	-0.0220
			(0.0138)	(0.0145)	(0.0155)
73-84 months post				-0.0280	-0.0285
				(0.0178)	(0.0184)
85-96 months post					-0.0150
					(0.0210)
Constant sample of cities	Yes	No	No	No	No
City pop. and GDP	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	0.392	0.457	0.457	0.457	0.457
R-squared	0.833	0.834	0.834	0.834	0.834
Number of events	26	39	39	39	39
N	8,312	12,169	12,169	12,169	12,169
Panel b.					
average post-period	-0.0206**	-0.0190**	-0.0188**	-0.0172**	-0.0181**
	(0.0086)	(0.0081)	(0.0084)	(0.0081)	(0.00800)

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

effect that lasts about five years.

B *Spatial scale of effect*

Subways overwhelmingly serve the areas close to the most central part of a city. Gonzalez-Navarro and Turner (2016) document that about 40% of all subway stations in existing subway systems

Table 6: Spatial decay

	0-10 km radius		10-25 km radius		25-50 km radius		50-150 km radius	
1-18 months post	-0.0205**	-0.0209**	-0.0167**	-0.0173**	-0.0115	-0.0121	-0.0045	-0.0049
	(0.0093)	(0.0095)	(0.0078)	(0.0081)	(0.0078)	(0.0082)	(0.0069)	(0.0071)
City pop. and GDP	No	Yes	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.457	0.457	0.400	0.400	0.366	0.366	0.295	0.295
R-squared	0.833	0.834	0.865	0.865	0.853	0.853	0.846	0.847
Number of events	39	39	39	39	39	39	39	39
N	12,169	12,169	12,169	12,169	12,535	12,535	11,897	11,897

*Note: Dependent variable is monthly mean AOD at a given distance from the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

lie within 5km of the center and another 30% within 10km of the center. Thus, we expect larger effects on AOD nearer to city centers than further away, particularly for this sample of new subway system cities.

Table 6 documents precisely this phenomenon. For reference, the first two columns of this table reproduce the first two columns of Table 3. Columns 3 and 4 are identical to columns 1 and 2, except that the dependent variable is mean monthly AOD between 10 and 25km of the center rather than within 10km. Columns 5 and 6 examine AOD between 25 and 50km, and columns 7 and 8 between 50 and 150km. As expected, the effect of subways on AOD decreases with distance from the city center and eventually goes to zero.

C *Further results*

Placebo test: Table A.5 presents the results of an important placebo test. To conduct this test, we match each subway city to the nearest city within 1000km that has a population within 20% of the target city. We are able to find such a match for 27 of our 39 subway cities. We replicate the regressions of Table 3 using placebo city outcomes in Table A.5. Simple t-tests fail to reject the hypothesis that the subway effect is different from zero in the placebo sample, but do reject the hypothesis that the subway effect is the same for placebo and subway cities.

A larger sample of cities and country-by-month indicators: We cannot include country-by-month indicators in our main specification because many of the countries in our sample experience only one subway opening during our study period. Moreover, our results are somewhat sensitive to the inclusion of city-by-calendar month indicators. Given this, we experiment with different

configurations of indicator variables to capture country and season effects to check the robustness of our findings.

First, our results are robust to using hemisphere-month fixed effects instead of year-by-continent and city-by-calendar month dummies. Second, in results presented in Appendix Table A.6, we extend our sample to include all 650 cities in the United Nation's World Urbanization Prospects data (DESA Population Division, 2014). This much larger sample of cities permits us to re-estimate and confirm our main results in specifications that use country-by-month indicators. This result indicates that our results hold even if we allow for completely flexible pollution trends at the country level.

Excluding observations with low pixel count: To check whether our main results are sensitive to the inclusion of city-months for which AOD coverage is sparce, we exclude observations with a number of pixel-days used to calculate AOD that falls below the tenth percentile of the distribution for a given satellite. The results are presented in Appendix Table A.7 and, as can be seen, they are very similar to those obtained in Table 3.

Expansions vs openings: Up until now, our investigations have focussed on the effects of the initial opening of a subway. We now turn to an investigation of subsequent expansions. To consider the effect of a city's first subway expansion (typically the second subway line in the system), as opposed to the system opening, we first restrict attention to the set of 14 cities for which we observe AOD for both the opening and the expansion. Column 1 of Table A.8 reports the results of a regression like the one reported in column 2 of Table 3 to estimate the effect on AOD of a city's first subway expansion over an 18 month treatment window. The point estimate of this effect is -0.015, smaller than the opening effect from Table 3, but we cannot distinguish this effect from zero.

Because we observe the whole history of each subway system, we can also identify the date of the first expansion for cities that do not open their subways during the period when we observe AOD. This increases the number of available expansions from 14 to 26. Column 3 of Table A.8 reports the estimate of the expansion effect on this larger set of cities. This estimate is smaller and more precisely estimated, and while we cannot distinguish it from zero, we also cannot distinguish it from -0.02. Columns 2 and 4 replicate the estimations of columns 1 and 3, but use any expansion as the treatment variable. These estimates are both close to zero and estimated with sufficient precision to allow us to distinguish them from -0.02, but not from zero. In sum, this table provides suggestive evidence that subway expansions are less important than openings.

Ridership: Table A.9 examines the effects of subway openings and expansions on subway ridership and confirms the conclusions suggested by Table A.8. An average subway opening results in ridership of about 62 people per 1000 of population. The first expansion, on average, increases

ridership by about 45 people per 1000 of population, and subsequent expansions have still smaller effects.

In Table A.2 we see that a subway system consists of 14.4 stations when it opens and that an average first expansion adds 15.7 stations. That is, an average expansion probably adds just about as much capacity as does an average opening. Together with our finding that first expansions seem to have a smaller effect than openings, this suggests that subway networks may not be subject to increasing returns to scale. Alternatively, it may be that subway expansions are built in places where there is less demand for subway travel than are the initial line.

Implications of longer pre-treatment period: The results presented so far are consistent with subway openings causing a decrease in AOD of about 0.02, and with this effect being about constant over at least the subsequent 5 to 8 years. If we instead consider a pre-period of around 48 months, then there are two noteworthy changes. First, from Table A.2 the set of cities for which we observe a 48 month pre and post-period begins with Gwanngju, South Korea and ends with Shenyang, China, so we are left with a dramatically smaller sample of 21 cities. Second, in this sample there are three extremely low pollution months about 24 months prior to subway opening. The three months do not appear to indicate a break in either the trend or the level of AOD. However, they are sufficiently different from neighboring months that they shift the 48 month pre-period AOD mean. With this said, in regressions analogous to those of Table 3 with the longer pre and post-treatment period and smaller sample, we estimate a decline in AOD around the time of subway opening which is only slightly smaller than what we report in Table 3. Figure A.3 illustrates this longer study period but is otherwise analogous to the panel (c) of Figure 5.

6. Subways, AOD and urban travel behavior

To reconcile our estimate of the effect of subway openings on AOD with other facts about traffic, subways, and pollution, consider the following back of the envelope calculations.

On the basis of travel survey data described in Akbar and Duranton (2017), residents of Bogota take about 2.69 trips per day and of these 19.3%, or 0.52 trips per day, are by private car or taxi. From Table A.2, twelve months after its subway opening an average city in our sample has a population of about 3.7m. Multiplying per capita trips by this population, if people drive at the same rate as does the population of Bogota then an average city in our sample generates about 2m trips by car or taxi per day. From Figure 1, in the second year of its operation an average subway in our data provides about 60 rides per 1000 of population. Applying this rate to a city of 3.7m, an average subway system provides about 200,000 rides per day. If all subway rides replace car or taxi rides, this is about 10% of all rides in our hypothetical city.

We can perform a similar calculation on the basis of US national averages using the 2009 National Household Transportation Survey.[12] The 2009 NHTS records that an average US household took 3.8 trips per day, about 90% by car (Duranton and Turner, 2017). If everyone in a city of 3.7m takes as many trips as an average American, then the city generates about 13m trips by car per day. If the subway system provides the same about 200,000 rides per day and if all subway rides replace car or taxi rides, then the subway opening should reduce car trips by about 1.5%.

For the sake of illustration, suppose a 1% reduction in traffic results in a 1% reduction in PM10 and that there is no demand response for car travel as drivers shift trips to the subway. In this case, the calculations above mean that a subway reduces PM10 by 10% in a city like Bogota and by 1.5% the US. These magnitudes bracket our sample average 4% AOD reduction. Thus, if a 1% reduction in traffic results in a 1% reduction in PM10 and that there is no demand response for car travel, then our estimated 4% AOD reduction seems reasonable.

With this said, the evidence for a direct relationship between driving and metropolitan PM10 levels is inconclusive. Friedman, Powell, Hutwagner, Graham, and Teague (2001) examine changes in traffic and changes in PM10 in Atlanta around the 1996 summer olympics. During this time, the city imposed restrictions on driving and saw traffic fall by about 2.8% over a 17 day period. During the same period, PM10 fell by 16.1%. That is, each 1% reduction in driving was associated with about a 5% reduction in PM10. In a related exercise, Gibson and Carnovale (2015) examine the effect of a change in Milan's congestion pricing program on both traffic and PM10. Their estimates allow us to calculate that a 4% reduction in traffic caused about a 1% reduction in PM10.[13] Finally, the US EPA attributes about 16% of US PM10 to on road vehicles.[14] Of these studies, only Friedman *et al.* (2001) finds that a one percent decrease in traffic is associated with at least a one percent decrease in PM10.

This has important implications for our calculations. Suppose that, in line with Gibson and Carnovale (2015), we require a 4% reduction in traffic for each 1% decrease in PM10. In this case, in order to achieve a 4% reduction in PM10, a subway opening would need to reduce traffic by 16%, not by 4%. This seems implausible. Alternatively, if subway openings reduce driving by 4% then we would expect them to reduce PM10 by 1%.

In addition, the evidence that drivers respond to increased road capacity seems compelling. Duranton and Turner (2011) find that metropolitan area traffic increases in direct proportion to road capacity, a finding that Hsu and Zhang (2014) confirm for Japan. While these papers study road expansions explicitly, the logic of their finding suggests that a reduction in traffic due to increased subway ridership ought to be met with almost exactly offsetting increases in the demand for automobile and truck travel. That is, that subways ought not to reduce traffic in urban areas.

[12] nhts.ornl.gov/2009/pub/stt.pdf, Table 3.

[13] From Gibson and Carnovale (2015) Table 6, treatment increased log PM10 by 0.0404 and, from Table 2, traffic by 26,725 cars. From Table 2, mean traffic and PM10 are 169,744 and 47.66. Using these values to calculate percentage changes we arrive at a 16% change in traffic and 4% change in PM10.

[14] U.S. EPA, Report on the Environment, `https://cfpub.epa.gov/roe/indicator.cfm?i=19`, accessed, April 2017.

Duranton and Turner (2011) corroborate this argument by looking at the relationship between the level of driving in a US metropolitan area and the stock of buses and subway cars in the MSA. Although the resulting estimates are imprecise, they do not contradict the hypothesis of no effect.

However, Duranton and Turner (2011) and Hsu and Zhang (2014) study, respectively, US and Japanese metropolitan areas, both with average populations of about 1m. In contrast, we consider a set of cities, all with subways, with average population of nearly 4m, of which only one is Japanese and two in the continental US. Thus, it may be that the demand responses estimated in Duranton and Turner (2011) and Hsu and Zhang (2014) do not extrapolate to our sample of larger and more international subway cities. Garcia-Lopez, Pasidis and Viladecans-Marsal (2017) provide suggestive evidence that the demand response to additional lanes is smaller for European cities with subways than without. With this said, the hypothesis of no demand response seems improbable.

This, too, has important implications for our calculations. Suppose that the demand response in our sample is only half what Duranton and Turner (2011) find for the US. In this case, in order to reduce driving by one trip, we would require two subway trips. This would halve the effect of subways on PM10 in our back of the envelope calculation.

To sum up, subway openings reduce metropolitan AOD levels by about 4% and probably carry approximately the same fraction of daily trips. However, a significant traffic demand response likely follows a subway opening, and on average a 1% traffic reduction probably reduces PM10 by less than 1%. It follows that the effect of subways on AOD probably does not operate entirely by diverting average trips to the subway. To cause the observed decrease in PM10 , subway openings must divert trips that are particularly polluting.

We can suggest two reasons why this might occur. First, marginal drivers who switch to the subway may be poor relative to the pool of all drivers, and if so, they may drive particularly dirty cars (or scooters). Thus, the availability of subways may affect pollution by affecting the composition of the stock of cars on the road. There does not seem to be a basis in the literature for estimating the magnitude of this effect. Second, evidence from hourly ridership indicates subways typically provide a disproportionate share of their trips at peak hours. In this case, subway trips may replace trips that occur at congested times and have high external congestion costs. If so, it may be that each car trips shifted to the subway has a disproportionate effect on pollution because they reduce pollution produced by other commuters. Anderson (2014) provides evidence, over the very short run, to support this idea.

7. Value of AOD reductions following subway openings

A *Value of health benefits from estimates in the economics literature*

Arceo, Hanna, and Oliva (2016) use data describing Mexico city between 1997 and 2006 to estimate a weekly infant death rate of 0.24 per 100,000 per $\mu g/m^3$ of PM10 . Thus, of 100,000 births, a 1

$\mu g/m^3$ decrease in ambient PM10 averts about 12.5 infant deaths.[15] Knittel, Miller, and Sanders (2016) use data from California between 2002 and 2007 to estimate a weekly infant death rate of 0.19 per 100,000 births per $\mu g/m^3$ of PM10. This estimate implies that of 100,000 births, a 1 $\mu g/m^3$ decrease in ambient PM10 averts 9.9 infant deaths. Chay and Greenstone (2003) consider data describing infant deaths in about 1000 US counties between 1978 and 1984 and estimate that a one unit $\mu g/m^3$ decrease in ambient TSP averts about 5.2 infant deaths per 100,000 births. Converting from TSP to PM10 is non-trivial, however PM10 $= 0.55 \times$ TSP is a sometimes used rule of thumb (World Bank Group and United Nations Industrial Development Organization, 1999). Rescaling the estimate from Chay and Greenstone (2003) implies that a one $\mu g/m^3$ decrease in ambient PM10 averts about 9.5 infant deaths per 100,000 births. In sum, these studies suggest that a one $\mu g/m^3$ decline in PM10 averts about 10 infant deaths per 100,000 births.[16]

That none of these estimates can be distinguished from the others despite a range of mean PM10 of from about 28 $\mu g/m^3$ for Knittel *et al.* (2016) to about 67 $\mu g/m^3$ for Arceo *et al.* (2016) suggests that the infant mortality response is approximately linear in PM10 (as Arceo *et al.* (2016) observe). Burnett *et al.* (2014) confirms the approximately linear dose-response relationship suggested by Arceo *et al.* (2016). More specifically, Burnett *et al.* (2014) surveys the large public health literature on the health consequences of PM10 and find that responses are approximately linear in the range from 5-100$\mu g/m^3$, although they find non-linearities outside of this range.

Figure 4(b) shows that most of our city-months fall in this 5-100$\mu g/m^3$ range. Therefore, in light of the results described above, we can reasonably assume a linear dose-response relationship in our sample.

Using these results, together with our estimates, allow us to estimate annual infant deaths averted by a subway opening for an average city. From Table 3, subway openings cause about a 0.02 unit decrease in AOD. Using column 2 of Table 2 to convert from AOD to PM10 gives about 2.0 $\mu g/m^3$ of PM10. At 10 infant deaths per 100,000 births per $\mu g/m^3$ of PM10, the number of averted deaths due to a subway opening in city i is given by

$$2.0 \times (10 \times 10^{-5}) \times \text{Birthrate}_i \times \text{Population}_i,$$

An average city in our sample has a population of about 3.7 million in the year before its subway opens. With a 2% birthrate, a subway opening in this city averts about 15 infant deaths per year.[17]

[15] An infant survives its first year if it survives 52 weeks. Thus a weekly death rate of 0.24×10^{-5} gives $[1 - (1 - (0.24 \times 10^{-5}))^{52}]10^5 = 12.47$ infant deaths per 100,000 births.

[16] Currie and Neidell (2005) use data describing PM10 and infant mortality in California between 1989 and 2000 and conclude that PM10 has no measurable effect on infant mortality. Knittel *et al.* (2016) and Arceo *et al.* (2016) both replicate the Currie and Neidell (2005) research design and find much smaller effects than the IV estimates reported above. We note that Jayachandran (2009) and Gutierrez (2010) also estimate the effects of particulates on infant mortality. We do not discuss their estimates because they do not present their results in a way that permits a conversion to mortality rates per $\mu g/m^3$ of PM10.

[17] The World Bank reports that the world average crude birth rate is 19.5/1000 (`http://data.worldbank.org/indicator/SP.DYN.CBRT.IN`, accessed April 2017).

With country-level birthrate data from the World Health Organization (2016c) and our population data, we can make this calculation somewhat more precisely. Specifically, imputing country level birth rates to cities, calculating the implied number of averted infant deaths for each city, and averaging over cities, we find that an average subway averts 9.4 infant deaths per year.

To monetize this benefit, we use country-adjusted values of a statistical life (VSL) to value averted infant deaths in each city.[18] Averaging over all cities, the value of averted infant deaths is \$21m per year. Our estimates do not allow us to conclude that subways continue to affect air quality beyond 5-6 years after their opening date. With a 5% discount rate, the present discounted value of this amount over five years is about 95.5m dollars. If the effect is permanent, the corresponding present value is 441m dollars.

B *Value of health benefits from the Global Burden of Disease methodology*

To extend our mortality calculations over the entire age distribution, we apply the methodology employed by Global Burden of Disease project (WHO, 2016b). This methodology is complicated and is described in detail in the Appendix Section A3.

We obtain integrated risk functions from Burnett *et al.* (2014) for five causes of mortality. These functions summarize the results of several epidemiological cohort studies, and consist of non-linear maps between PM2.5 concentrations and mortality risk ratios. We quantify the contribution of air pollution to age and disease-specific mortality by computing the *population attributable fraction*. This is the percentage mortality reduction that would occur if PM2.5 concentrations were reduced to a counterfactual exposure level. We proceed in three steps. First, using coefficients in column 5 of Table 2 (Terra), we predict city-level PM2.5 concentrations during the 12 months preceding the opening of a subway. Second, applying the integrated risk functions, we calculate the population attributable fractions associated with a 1.102 $\mu g/m^3$ PM2.5 decrease from the predicted pre-subway exposure level.[19] Third, using city population and country-level death rates from the World Health Organisation (WHO, 2016c), we construct city-specific mortality for each disease and age class. The total number of avoided deaths in each city is obtained by applying the population attributable fractions to the mortality rates, summing over every disease. Averaging over all cities, the 0.02 unit decrease in AOD that follows a subway opening saves about 222 lives per year.

Valuing these lives, as before, with country-adjusted values of a statistical life, the value of averted deaths is \$594m per year. With a 5% discount rate, the present discounted value of this amount over five years is about 2.7b dollars. If the effect is permanent, the corresponding present value is 12.5b dollars.

[18] Specifically, we take Viscusi and Aldy's (2003) 0.6 elasticity of VSL with respect to income, and impute a country's VSL from the U.S. value of \$6m.

[19] From column 5 of Table 2, a 0.02 unit decrease in AOD converts to a $0.02 * 55.12 = 1.102\ \mu g/m^3$ decrease in PM2.5.

C *Discussion*

These benefit calculations are obviously crude. They do not account for morbidity, for possible effects on labor productivity, nor for the fact that subways may reduce pollutants other than particulates. While the magnitudes of these effects remain uncertain, they are almost surely positive, and possibly large, e.g., Chang, Zivin, Gross, and Neidell (2016) or Murray (2016). Thus, we might reasonably expect that a complete accounting for the health and productivity related benefits of subway induced improvements in air quality would lead to a much larger value than we describe above.

The mean length of track for a newly opened subway system in our study period is 19.2km. Baum-Snow and Kahn (2005) examine 16 US subway systems and estimate constructions costs ranging between 25m and 287m dollars per mile. Using this range of cost estimates, the cost of construction for an average subway system in our sample ranges from 298m to about 3.4b dollars. Our estimates of the present value of avoided infant mortality in an average subway city range between 95.5m and 441m, depending on whether the subway effect on pollution lasts for five years or is permanent. Our estimates of the present value of avoided all-age mortality in an average subway city range between 2.7b and 12.5b, again depending on whether the subway effect on pollution lasts for five years or is permanent. Comparing these magnitudes suggests that the value of subway induced improvements to air quality may be a substantial fraction of subway construction costs.

8. Conclusion

Column 2 of Table 3 indicates that AOD fell by about 0.021 in the 10k disk surrounding a city center during the 18 months following an average subway opening. This effect is robust to econometric specification, and is estimated precisely, the 95% CI is [-0.040,-0.002]. Mean monthly AOD in the regression sample before the introduction of subways is 0.49. Dividing, the mean AOD reduction from a subway opening is about 4% with a 95% confidence interval ranging from 0.5% to 8%.

From Table 1, AOD readings from Terra in 2000 and 2014 are 0.38 and 0.45. Our 0.02 estimated subway effect is only about 4% of the level of AOD, but it is more than 25% of the 14 year change. Comparing across continents in 2014, we see that the difference between Europe and North America is about 0.07, three and a half subway effects, and between Europe and Asia the difference is 0.38, about 20 subway effects. In all, this suggests that subways may play a moderately important role in determining AOD in a city.

To the extent that our data allows us to consider longer post-opening periods, this effect of subways on AOD appears to be approximately constant during the 5-6 years after an opening, and indeterminant beyond this time. As we consider larger areas around the city center, the effects of subway openings attenuate in intuitive ways.

The effects of subway expansions seem to be smaller than those of subway openings, an effect that we observe both in changes in AOD and ridership. To the extent that public policy encourages subway construction, this suggests that openings are relatively more important than expansions. We note that the decreasing marginal effect of subway expansion seems to be broadly consistent with the slight decreasing returns in the effects of metropolitan road networks observed in Couture, Duranton, and Turner (2016).

Extant estimates of the effects of particulates on mortality suggest that they are sufficiently poisonous that the small nominal reductions from subway openings are economically important. In particular, they appear to be large enough to justify moderate subsidies for subway construction in environments where more direct policy instruments for managing automobile pollution are not available.

On the other hand, the channel through which subways affect pollution remains somewhat uncertain. Our calculation suggests that observed levels of ridership account for a significant fraction of metropolitan travel. In order for this level of ridership to cause the observed 4% decline in AOD, it seems likely that trips diverted to the subway are particularly polluting, either because they use particularly polluting vehicles or because they occur at especially congested times.

References

Akbar, Prottoy and Gilles Duranton. 2017. Measuring the cost of congestion in a highly congested city: Bogota. Processed, University of Pennsylvania.

Anderson, Michael L. 2014. Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *The American Economic Review* 104(9): 2763–2796.

Andrews, Donald W.K. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society* : 821–856.

Andrews, Donald W.K. 2003. Tests for parameter instability and structural change with unknown change point: A corrigendum. *Econometrica* : 395–397.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from Mexico City. *The Economic Journal* 126(591): 257–280.

Baum-Snow, Nathaniel and Matthew E. Kahn. 2005. Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000. *Brookings-Wharton Papers on Urban Affairs: 2005* 1(4): 147–197.

Billings, Stephen B. 2011. Estimating the value of a new transit option. *Regional Science and Urban Economics* 41(6): 525–536.

Brauer, Michael, Greg Freedman, Joseph Frostad, Aaron Van Donkelaar, Randall V. Martin, Frank Dentener, Rita van Dingenen, Kara Estep, Heresh Amini, and Joshua S. Apte. 2015. Ambient air pollution exposure estimation for the global burden of disease 2013. *Environmental Science & Technology* 50(1): 79–88.

Burnett, Richard T *et al.* 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environmental Health Perspectives* 122(4): 397.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics* 90(3): 414–427.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy* 8(3): 141–69.

Chay, Kenneth and Michael Greenstone. 2003. The impact of air pollution on infant mortality. *Quarterly Journal of Economics* 118(3): 1121–1167.

Chen, Yihsu and Alexander Whalley. 2012. Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal* 4(1): 58–97.

Couture, Victor, Gilles Duranton, and Matthew A. Turner. 2016. Speed. Processed, Brown University.

Currie, Janet and Matthew Neidell. 2005. Air pollution and infant health: What can we learn from California's recent experience? *Quarterly Journal of Economics* 120(3): 1003–1030.

DESA Population Division, UN. 2014. World urbanization prospects: the 2014 revision, cd-rom edition.

Duranton, Gilles and Matthew A. Turner. 2011. The fundamental law of road congestion: Evidence from US cities. *American Economic Review* 101(6): 2616–2652.

Duranton, Gilles and Matthew A. Turner. 2017. Urban form and driving. Processed, Brown University.

Feenstra, Robert, Robert Inklaar, and Marcel Timmer. 2015. The next generation of the penn world table. *American Economic Review* 105(10): 3150–3182. Available for download at www.ggdc.net/pwt
Accessed: 2016-09-06.

Friedman, Michael S., Kenneth E. Powell, Lori Hutwagner, LeRoy M. Graham, and W. Gerald Teague. 2001. Impact of changes in transportation and commuting behaviors during the 1996 summer olympic games in atlanta on air quality and childhood asthma. *JAMA* 285(7): 897–905.

Garcia-Lopez, Miquel-Angel, Ilias Pasidis, and Elisabet Viladecans-Marsal. 2017. Highway congestion and air pollution in Europe's cities. *Working paper* .

Gibbons, Stephen and Stephen Machin. 2005. Valuing rail access using transport innovations. *Journal of Urban Economics* 57(1): 148–1698.

Gibson, Matthew and Maria Carnovale. 2015. The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics* 89: 62–73.

Gonzalez-Navarro, Marco and Matthew A. Turner. 2016. Subways and urban growth: Evidence from earth. Processed, Brown University.

Gupta, Pawan, Sundar A. Christopher, Jun Wang, Robert Gehrig, Yc Lee, and Naresh Kumar. 2006. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment* 40: 5880–5892.

Gutierrez, Emilio. 2010. Using satellite imagery to measure the relationship between air quality and infant mortality: an empirical study for Mexico. *Population and Environment* 31: 203–222.

Hansen, Bruce E. 2000. Testing for structural change in conditional models. *Journal of Econometrics* 97(1): 93–115.

Harris, I., P.D. Jones, T.J. Osborn, and D.H. Lister. 2014. Updated high-resolution grids of monthly climatic observations - the cru ts3.10 dataset. International Journal of Climatology, 34(3), 623-642. Available for download at https://crudata.uea.ac.uk/cru/data/hrg/
Accessed: 2015-11-20.

Hsu, Wen-Tai and Hongliang Zhang. 2014. The fundamental law of highway congestion: Evidence from national expressways in Japan. *Journal of Urban Economics* 81: 65–76.

Jacob, Daniel. 1999. *Introduction to atmospheric chemistry*. Princeton University Press.

Jayachandran, Seema. 2009. Air quality and early-life mortality evidence from Indonesia's wildfires. *Journal of Human resources* 44(4): 916–954.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. Caution, drivers! children present: Traffic, pollution, and infant health. *Review of Economics and Statistics* 98(2): 350–366.

Kumar, Naresh, Allen Chu, and Andrew Foster. 2007. An empirical relationship between pm2.5 and aerosol optical depth in Delhi metropolitan. *Atmospheric Environment* 41: 4492–4503.

Kumar, Naresh, Allen D. Chu, Andrew D. Foster, Thomas Peters, and Robert Willis. 2011. Satellite remote sensing for developing time and space resolved estimates of ambient particulate in Cleveland, OH. *Aerosol Science and Technology* 45: 1090–1108.

Levy, R. C., S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, F. Patadia, and N. C. Hsu. 2013. The collection 6 modis aerosol products over land and ocean. *Atmospheric Measurement Techniques* 6(11): 2989–3034. URL `http://www.atmos-meas-tech.net/6/2989/2013/`.

Levy, Rob and Christina Hsu et al. 2015*a*. MODIS Terra l2 aerosol product, nasa modis adaptive processing system, goddard space flight center, usa. http://dx.doi.org/10.5067/MODIS/MOD04_L2.006. Accessed: 2016-01-15.

Levy, Rob and Christina Hsu et al. 2015*b*. MODIS Aqua l2 aerosol product, nasa modis adaptive processing system, goddard space flight center, usa. http://dx.doi.org/10.5067/MODIS/MYD04_L2.006. Accessed: 2016-01-15.

Murray, Christopher J. L. 2016. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015. *Lancet* 388(October 8): 1659–1724.

Remer, L. A., S. Mattooand R. C. Levy, and L. A. Munchak. 2013. Modis 3km aerosol product: algorithm and global perspective. *Atmospheric Measurement Techniques* 6: 1829–1844.

Viscusi, W. Kip and Joseph E. Aldy. 2003. The value of a statistical life: a critical review of market estimates throughout the world. *Journal of risk and uncertainty* 27(1): 5–76.

Voith, Richard. 1997. Fares, service levels, and demographics: What determines commuter rail ridership in the long run? *Journal of Urban Economics* 41(2): 176–197.

Vollmer Associates, LLP, SYSTRA Consulting, Inc., and Allee King Rosen & Fleming, Inc. 2011. Manhattan east side transit alternatives (mesa)/second avenue subway summary report. MTA New York City Transit.

World Bank Group and United Nations Industrial Development Organization. 1999. *Pollution prevention and abatement handbook, 1998: toward cleaner production*. World Bank Publications.

World Health Organization. 2006. WHO air quality guidelines. Geneva: World Health Organization.

World Health Organization. 2016a. Global urban ambient air pollution database. http://www.who.int/phe/healt_topics/outdoorair/databases/cities/en/. Accessed: 2017-04-02.

World Health Organization. 2016b. Ambient air pollution: A global assesment of exposure and burden of disease.

World Health Organization. 2016c. Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000-2015.

Appendix

A1 Ridership data

We gathered subway ridership data (unlinked trips) for 30 of the subway systems in our sample, mostly from annual reports or statistical agencies. In 13 cases we were either not able to find data on ridership at all, the data were not available from the opening date, or the ridership data was aggregated across cities or other rail systems. Data sources for each of the cities we were able to obtain usable data are detailed in Table 1. For ten of the cities we were able to obtain data, ridership was reported at the monthly level. For the other 20, quarterly or yearly data was available and we used linear interpolation to create a monthly level ridership dataset.

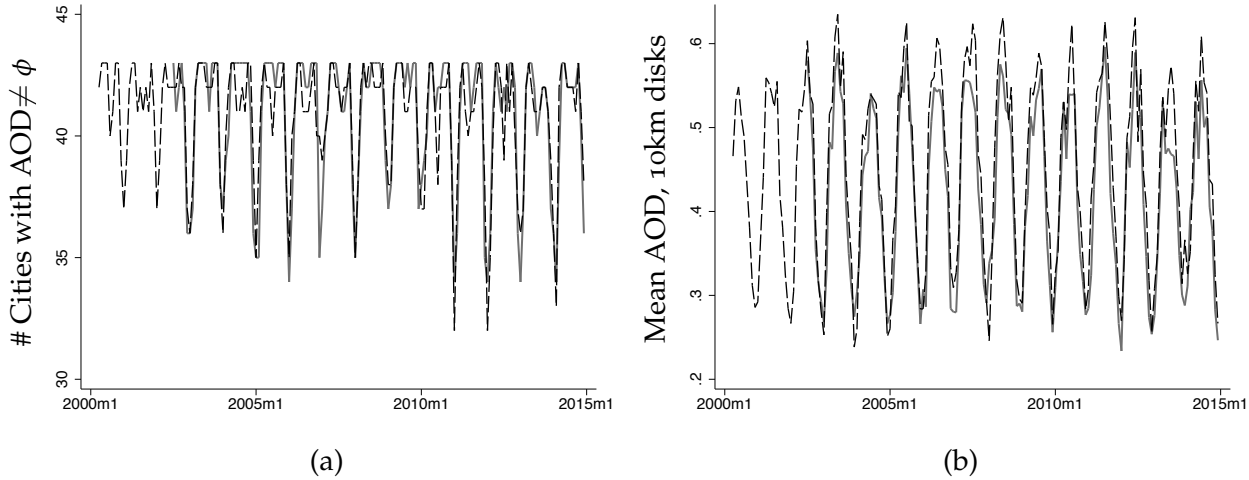
Table A.1: Ridership data sources

City	Source
Almaty (Kazakhstan)	International Metro Association Reports
Bangalore (India)	Bangalore Metro Operational Performance
Brescia (Italy)	Brescia Mobilitá Reports
Copenhagen (Denmark)	Statistics Denmark
Daejon (South Korea)	Daejon Metropolitan Rapid Transit Corporation
Delhi (India)	Delhi Metro Rail Corporation Annual Reports
Dubai (UAE)	Dubai Road and Transport Auth.: Annual Statistical Reports
Gwangju (South Korea)	Gwangju Subway Reports
Hangzhou (China)	Hangzhou Statistical Yearbook
Istanbul (Turkey)	Metro Istanbul Statistics (Only line M2)
Kazan (Russia)	International Metro Association Reports
Kaohsiung (Taiwan)	Kaohsiung Rapid Transit Corp. Transport Volume Statistics
Las Vegas (USA)	NTA National Transit Database and Webarchive lvmonorail
Lausanne (Switzerland)	Transports Lausanne Annual Reports
Lima (Peru)	Ministerio de Transportes y Comunicaciones Perú
Mashhad (Iran)	Mashhad Urban Railway Corp.: Planning and Development
Naha (Japan)	Japan Ministry of Land, Infrastructure, Transport and Tourism
Palma (Spain)	Instituto Nacional de Estadística España
Porto (Portugal)	Statistics Portugal, Light Rail (Metro) Survey
San Juan Puerto Rico (USA)	Instituto de Estadísticas de Puerto Rico
Santo Domingo (DR)	Oficina para el Reordenamiento del Transporte
Seattle (USA)	Sound Transit (Only Central Link Line) Performance Reports
Seville (Spain)	Instituto Nacional de Estadística España
Shenzhen (China)	Shenzhen Municipal Transportation Commission
Shenyang (China)	Shenyang Statistical Information Net
Suzhou, Jiangsu (China)	Suzhou Statistical Yearbook
Tehran (Iran)	Tehran Metro Research and Development
Turin (Italy)	Gruppo Torinese Transporti Reports
Valparaiso (Chile)	Memoria Anual Metro Valparaiso
Xian, Shaanxi (China)	Xian Bureau of Statistics

We were not able to obtain ridership data from the time of opening for the following 13 cities in the sample: Algiers (Algeria), Brasilia (Brazil), Bursa (Turkey), Chengdu (China), Chongqing (China), Dalian (China), Izmir (Turkey), Kunming (China), Maracaibo (Venezuela), Nanjing (China), Rennes (France), Valencia (Venezuela), and Wuhan (China).

A2 AOD data

Figure A.1: MODIS Terra and Aqua AOD data



(a) (b)

Panel (a) gives count of subway cities months with AOD 10km measurements by month for Terra (dashed black) and Aqua (gray). Panel (b) shows mean AOD within 10km of the center of subway cities, averaged over cities, by month for Terra (dashed black) and Aqua (gray).

The Moderate Resolution Imaging Spectroradiometers (MODIS) aboard the Terra and Aqua Earth observing satellites measure the ambient aerosol optical depth (AOD) of the atmosphere almost globally. We use MODIS Level-2 daily AOD products from Terra for February 2000-December 2014 and Aqua for July 2002-December 2014 to construct monthly average AOD levels in cities. We download all the files from the NASA File Transfer Protocol.[20]

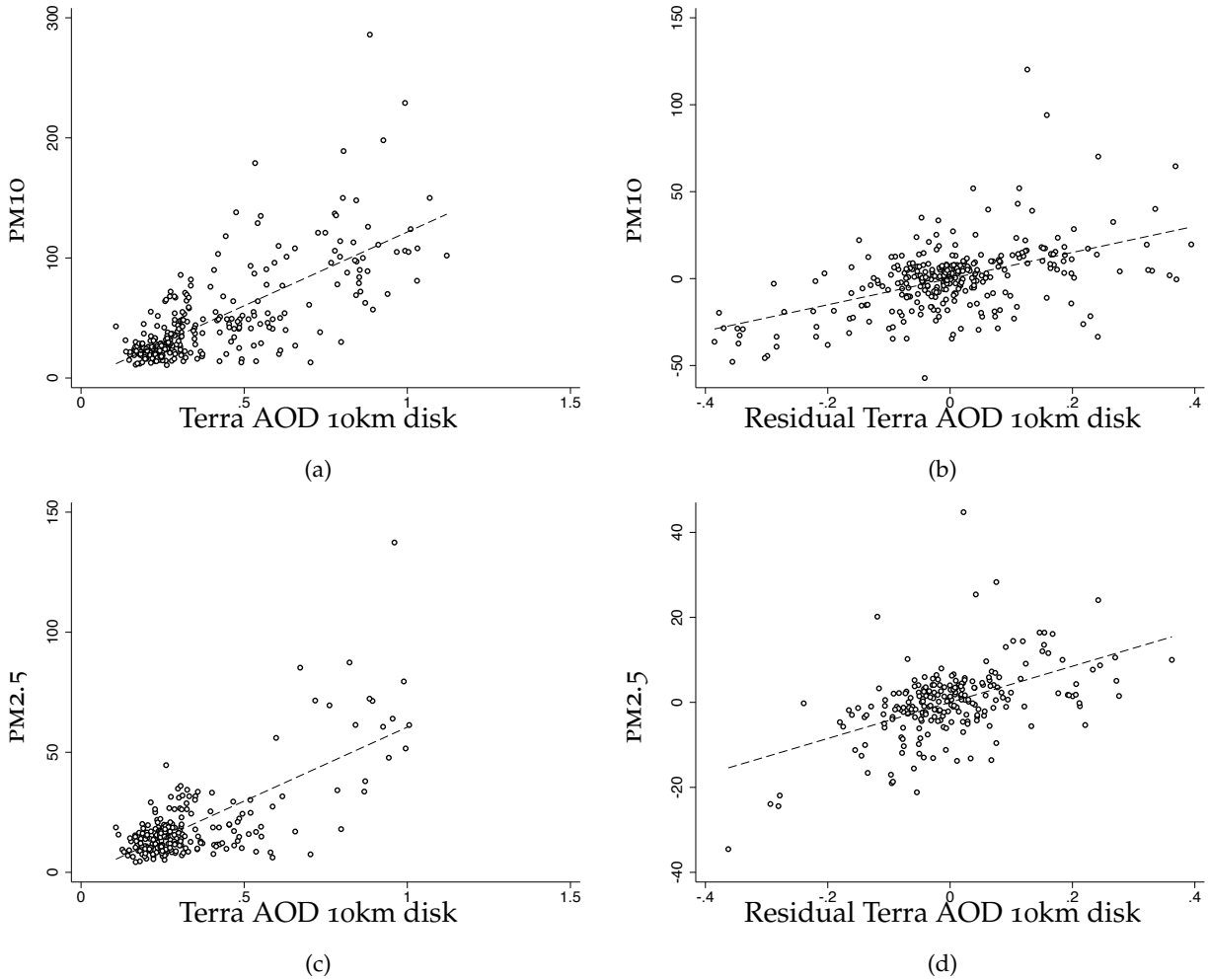
There are four MODIS Aerosol data product files: $MOD04\_L2$ and $MOD04\_3K$, containing data collected from the Terra platform; and $MYD04\_L2$ and $MYD04\_3K$, containing data collected from the Aqua platform. We use products $MOD04\_L2$ and $MYD04\_L2$ to get AOD measures at a spatial resolution (pixel size) of approximately 10 x 10 kilometers and products $MOD04\_3K$ and $MYD04\_3K$ to get AOD measures at a spatial resolution of approximately 3 x 3 kilometers. Each product file covers a five-minute time interval based on the start time of each MODIS granule. The product files are stored in Hierarchical Data Format (HDF) and we use the "Optical Depth Land And Ocean" layer, which is stored as a Scientific Data Set (SDS) within the HDF file, as our measure of aerosol optical depth. The "Optical Depth Land And Ocean" dataset contains only the AOD retrievals of high quality.

We convert all HDF formatted granules to GIS compatible formats using the HDF-EOS To GeoTIFF Conversion Tool (HEG) provided by NASA's Earth Observing System Program[21]. We consolidate every GeoTIFF granules into a global raster for each day using ArcGIS. First, we keep only AOD values that do contain information. The missing value is -9999 in AOD retrievals.

[20] ftp://ladsweb.nascom.nasa.gov/allData/6/

[21] The most recent version of the software, HEG Stand-alone v2.13, can be downloaded at http://newsroom.gsfc.nasa.gov/sdptoolkit/HEG/HEGDownload.html

Figure A.2: Plots of ground-based PM10 and PM2.5 vs. MODIS AOD



Note: Panel (a) Plot of ground-based PM10 against Terra MODIS AOD in a 10km disk. Panel (b) Plot of ground-based PM10 residual against Terra MODIS AOD in a 10km disk residual. Panel (c) Plot of ground-based PM2.5 against Terra MODIS AOD in a 10km disk. Panel (d) Plot of ground-based PM2.5 residual against Terra MODIS AOD in a 10km disk residual. NB: Scales not constant across graphs.

Second, we create a raster catalog with all the granules for a given day and calculate the average AOD value using the Raster Catalog to Raster Dataset tool.

Figure A.1 provides more information about the coverage of the two satellites and the prevalence of missing data. The black dashed line in panel (a) of figure gives the count of cities in our data for which we calculate an AOD from the Terra satellite reading for each month of our study period. These are cities for which there is at least one pixel within 10km of the center on one day during the relevant month. Since most of the cities in our data are in the Northern hemisphere, we see a strong seasonal pattern this series. The light gray line in this figure reports the corresponding quantity calculated from the Aqua satellite reading. Since Aqua became operational after Terra, the Aqua series begins later. The Aqua satellite data tracks the Terra data closely, but at a slightly lower level. Panel (b) of figure A.1 reports city mean AOD data for all city-months in our sample over

the course of our study period. As for the other series, this one too exhibits seasonality, although this will partly reflect a composition effect. As we see in panel (a) not all cities are in the data for all months. As in the first two panels, the dark line describes AOD readings from Terra and the light gray, Aqua.

Finally, figure A.2 shows the relationship between the ground based measurements and MODIS AOD. Panel (a) of this figure plots ground based PM10 against Terra AOD in a 10km disk. That is, the raw data on which column 1 of table 2 is based. We see a strong positive relationship. Panel (b) shows a plot of the residuals of the regression in column 3 of table 2 against the residuals of a regression of 10km Terra PM10 on the control variables used in the same regression. Again, we see a strong positive slope. Note that the scales on the two graphs are not the same. The bottom two panels are the same as the top, but are based on ground based PM2.5 measures. Again we see a clear positive slope in both plots.

A3 Global Burden of Disease based mortality estimates

The integrated risk functions in Burnett *et al.* (2014) express the likelihood of dying from a disease at current PM2.5 exposure, relative to an environment where PM2.5 concentrations are set to a baseline harmless level of exposure. If D_d is the event of dying from disease d, the risk ratio (RR) of being exposed to PM2.5 concentration c is given by $RR_d(c, \bar{c}) = P(D_d \mid c)/P(D_d \mid \bar{c})$, where $\bar{c}$ denotes the baseline harmless concentration. Burnett *et al.* (2014) model $RR_d(c, \bar{c})$ to exhibit diminishing marginal risk: $RR(c,\bar{c}) = 1 + \alpha(1 - e^{-\gamma(c-\bar{c})^{\delta}})$ if $c > \bar{c}$, and $RR(c,\bar{c}) = 1$ otherwise, with $\bar{c}$ assumed to lie uniformly between 5.8 and 8.8μg/m^3. We refer the reader to Burnett *et al.* (2014) for details regarding the parametrization and estimation of these functions for each disease.

As described in the main text, we obtain RR functions for five diseases: ischemic heart disease, cerebrovascular disease (stroke), chronic obstructive pulmonary disease, lung cancer, and lower respiratory infection. For deaths attributable to stroke and ischemic heart disease, the integrated risk functions are age-specific. To construct population attributable fractions (PAF) for every disease and, when applicable, every age-group, we first predict pre and post-subway PM2.5 concentrations using the regression specification in column 5 of table 2. Specifically, we obtain predicted PM2.5 values from the annual city average of AOD (and all other covariates) during the 12 months preceding the subway opening. The post-subway PM2.5 concentrations are obtained by subtracting $0.02 \times 55.12 = 1.102$ μg/m^3 to the pre-subway concentration, where 0.02 is the subway AOD effect from table 3, and 55.12 is the AOD coefficient in column 5 of table 2.

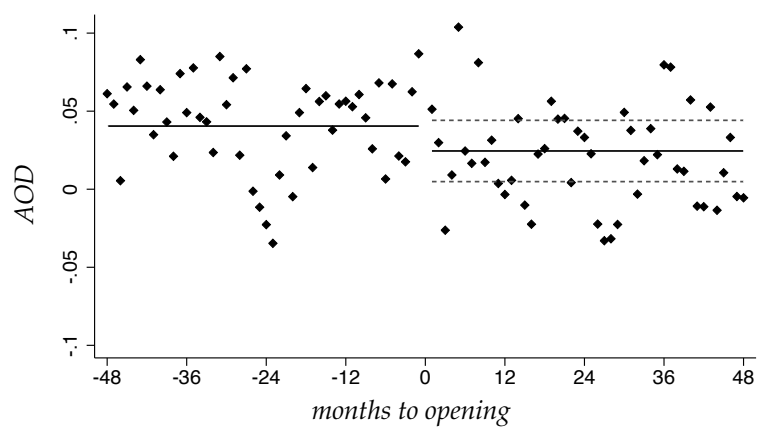
Let c_0 and c_1 respectively denote the pre and post-subway PM2.5 concentrations in a given city. For the purpose of our burden of disease calculations, the relevant risk ratio is $RR_d(c_1,c_0) = P(D_d \mid c_1)/P(D_d \mid c_0)$. Using the $RR_d(c,\bar{c})$ functions in Burnett *et al.* (2014), we obtain this number by computing $RR_d(c_1,c_0) = RR_d(c_1,\bar{c})/RR_d(c_0,\bar{c})$. Here, $RR_d(c_1,c_0)$ expresses how much less likely it is that individuals die of disease d when exposed to concentration c_1, relative to concentration c_0. Assuming that 100% of the city population is exposed to c_0 and then c_1, the population attributable

fraction is then just $PAF_d = 1 - RR_d(c_1,c_0) = 1 - P(D_d \mid c_1)/P(D_d \mid c_0)$. Interpreting $P(D_d \mid c)$ as the fraction of the total population that died of disease d when exposed to PM2.5 concentration c, we find that PAF_d represents the fraction of total deaths from d that occurred because of incremental pollution $c_0 - c_1$.

Finally, for each city i, we calculate the number of death attributable to disease d in age-group a (denoted M_{ida} below). We use disease-specific country-level death rates from the World Health Organisation (WHO, 2016c) and apply them to city populations. Mortality data from the WHO is only available in 2000, 2005, 2010, and 2015. We use the year closest to a city's subway opening year. The total number of avoided death in city i is given by $\sum_d \sum_a PAF_{ida} \cdot M_{ida}$.

A4 Appendix Tables and Figures

Figure A.3: AOD during the 48 months before and after subway openings



(c) Event study

Notes: Event study during 48 months before and after subway openings, constant sample of 21 cities.

Table A.2: City level descriptive statistics

City	Plan approved	Construction begins	Opening date	Track km	Stations added			Daily ridership	Mean AOD	SD AOD	City population	Country GDP PC
					opening	1st exp.	2nd exp.					
Tehran (Iran)	1970	Jan. 1977	Feb. 2000	10.2	8	10	1	290,740	0.57	0.18	7,145	7,602
Izmir (Turkey)	1992	Mar. 1995	Aug. 2000	11.1	10	n.a.	n.a.		0.30	0.09	2,248	11,264
Istanbul (Turkey)	1987	Sep. 1992	Sep. 2000	6.4	5	1	15	64,360	0.32	0.13	8,963	11,221
Brasilia (Brazil)	1991	Jan. 1992	Mar. 2001	42.7	15	n.a.	n.a.		0.11	0.07	3,024	8,445
Rennes (France)	1989	Jan. 1997	Mar. 2002	10.3	15	n.a.	n.a.		0.22	0.11	283	33,534
Bursa (Turkey)	1997	Jan. 1998	Aug. 2002	19.1	17	n.a.	n.a.		0.27	0.09	1,290	10,544
Copenhagen (Denmark)	1995	Nov. 1996	Nov. 2002	20.2	13	n.a.	n.a.	69,576	0.19	0.10	1,100	35,935
Porto (Portugal)	1996	Mar. 1999	Dec. 2002	6.2	7	4	5	21,951	0.23	0.09	1,266	22,919
Delhi (India)	1995	Jan. 1998	Dec. 2002	12.7	10	3	22	117,566	0.88	0.35	17,504	2,325
Dalian (China)	1999	Sep. 2000	May 2003	52.4	12	6	n.a.		0.59	0.27	3,157	5,473
Naha (Japan)	1996	Nov. 1996	Aug. 2003	11.6	15	n.a.	n.a.	319,033	0.30	0.15	304	34,085
Gwangju (South Korea)	1994	Aug. 1996	Apr. 2004	10.4	13	n.a.	n.a.	34,639	0.46	0.23	1,400	27,562
Las Vegas (USA)	1993	Aug. 2001	Jul. 2004	4.3	7	n.a.	n.a.	29,078	0.51	0.20	1,565	50,048
Wuhan (China)	1999	Dec. 2012	Sep. 2004	28.1	25	20	n.a.		1.01	0.31	7,110	6,249
Shenzhen (China)	1992	Dec. 1998	Dec. 2004	41.0	18	69	n.a.	204,944	0.85	0.27	8,409	6,400
San Juan Puerto Rico (USA)	1992	Jul. 1996	Apr. 2005	17.1	16	n.a.	n.a.	24,000	0.22	0.11	2,492	50,799
Kazan (Russia)	1989	Aug. 1997	Aug. 2005	9.4	5	n.a.	n.a.	46,388	0.24	0.16	1,121	15,304
Nanjing (China)	1994	Dec. 2000	Aug. 2005	27.3	16	38	n.a.		0.83	0.31	5,263	6,883
Valparaiso (Chile)	1999	May 2002	Nov. 2005	61.4	20	n.a.	n.a.	28,053	0.11	0.03	843	15,561
Turin (Italy)	1999	Dec. 2000	Feb. 2006	9.5	11	n.a.	n.a.	23,878	0.31	0.15	1,713	34,143
Daejon (South Korea)	1996	Jan. 1996	Mar. 2006	10.6	12	n.a.	n.a.	25,095	0.42	0.25	1,449	29,209
Valencia (Venezuela)	1994	Nov. 1997	Nov. 2006	3.9	3	n.a.	n.a.		0.26	0.12	1,555	13,049
Maracaibo (Venezuela)	1998	Mar. 2004	Nov. 2006	1.4	2	n.a.	n.a.		0.41	0.17	1,945	13,049
Kaohsiung (Taiwan)	1994	Jan. 2001	Mar. 2008	37.3	36	n.a.	n.a.	105,618	0.58	0.24	1,510	36,785

Continued on next page

Table A.2 – continued from previous page

City	Plan approved	Construction begins	Opening date	Track km	Stations added opening	Stations added 1st exp.	Stations added 2nd exp.	Daily ridership	Mean AOD	SD AOD	City population	Country GDP PC
Palma (Spain)	2004	Aug. 2005	Jul. 2008	7.1	9	6	n.a.	4,348	0.19	0.07	381	34,601
Lausanne (Switzerland)	2000	Feb. 2004	Oct. 2008	5.5	14	n.a.	n.a.	62,638	0.21	0.07	334	51,771
Santo Domingo (DR)	2005	Nov. 2005	Jan. 2009	11.3	16	13	n.a.	44,946	0.35	0.19	2,542	10,537
Seville (Spain)	1999	Aug. 2005	Apr. 2009	17.6	21	n.a.	n.a.	46,922	0.21	0.08	695	33,882
Seattle (USA)	1996	Nov. 2003	Jul. 2009	14.6	8	n.a.	n.a.	20,068	0.17	0.08	3,052	49,281
Dubai (UAE)	2005	Mar. 2006	Sep. 2009	40.7	21	16	n.a.	123,980	0.52	0.24	1,748	58,319
Chengdu (China)	2000	Dec. 2005	Sep. 2010	15.6	16	20	n.a.		1.01	0.33	6,420	10,017
Shenyang (China)	1999	Nov. 2005	Sep. 2010	31.2	22	18	n.a.	279,395	0.55	0.29	5,770	10,017
Chongqing (China)	1998	Jun. 2007	Jul. 2011	15.2	13	10	n.a.		0.97	0.24	11,880	10,636
Lima (Peru)	1986	Mar. 2010	Jul. 2011	19.0	16	n.a.	n.a.	82,751	0.76	0.23	9,251	10,506
Xian, Shaanxi (China)	1994	Sep. 2006	Sep. 2011	17.0	16	n.a.	n.a.	193,324	0.83	0.29	5,452	10,759
Bangalore (India)	2003	Apr. 2007	Oct. 2011	5.6	6	n.a.	n.a.	18,433	0.40	0.17	8,908	4,698
Mashhad (Iran)	1994	Dec. 1999	Oct. 2011	20.9	21	n.a.	n.a.	75,330	0.53	0.17	2,785	15,876
Algiers (Algeria)	1988	Mar. 1993	Nov. 2011	8.9	10	n.a.	n.a.		0.25	0.12	2,492	13,193
Almaty (Kazakhstan)	1980	Sep. 1988	Dec. 2011	8.0	7	n.a.	n.a.	16,944	0.32	0.08	1,454	22,010
Suzhou, Jiangsu (China)	2002	Dec. 2007	Apr. 2012	26.6	24	1	n.a.	126,335	0.87	0.27	4,639	11,187
Kunming (China)	2009	May 2010	Jun. 2012	38.9	14	n.a.	n.a.		0.35	0.27	3,571	11,308
Hangzhou (China)	2005	Mar. 2007	Nov. 2012	45.9	31	n.a.	n.a.	305,314	0.85	0.23	5,821	11,612
Brescia (Italy)	2000	Jan. 2004	Mar. 2013	14.3	17	n.a.	n.a.	40,719	0.32	0.15	456	35,912
Average	1995	Jul. 2000	Feb. 2007	19.2	14.4	15.7	10.8	97,514	0.46	0.18	3,719	20,011

Note: Stations and daily ridership reported 12 months after opening. Mean and SD AOD columns report mean and standard deviation values in a 10km radius circle using Terra satellite monthly observations from 2000-2014. Metropolitan area population in thousands.

Table A.3: Subway opening and AOD by 6 month period, pre- and post-system opening

	(1)	(2)	(3)	(4)	(5)	(6)
1-6 months post	-0.0134	-0.0140	-0.0139	-0.0145	-0.0144	-0.0149
	(0.0134)	(0.0136)	(0.0135)	(0.0136)	(0.0130)	(0.0133)
7-12 months post	-0.0249*	-0.0249*	-0.0250*	-0.0254*	-0.0232*	-0.0241*
	(0.0137)	(0.0138)	(0.0135)	(0.0133)	(0.0127)	(0.0129)
13-18 months post	-0.0226	-0.0226	-0.0225	-0.0231	-0.0189	-0.0201
	(0.0157)	(0.0162)	(0.0159)	(0.0159)	(0.0155)	(0.0159)
7-12 months pre	0.00668	0.00652	0.00569	0.00615	0.00892	0.00926
	(0.0120)	(0.0121)	(0.0122)	(0.0121)	(0.0127)	(0.0127)
13-18 months pre	-0.00711	-0.00654	-0.00793	-0.00707	-0.00140	-0.000921
	(0.0113)	(0.0112)	(0.0113)	(0.0112)	(0.0118)	(0.0120)
city pop.& GDP	No	Yes	No	Yes	No	Yes
city-level trends	No	No	Yes	Yes	No	No
city-level pre/post trends	No	No	No	No	Yes	Yes
Mean dep. var.	0.457	0.457	0.457	0.457	0.457	0.457
R-squared	0.833	0.834	0.838	0.838	0.839	0.839
Number of events	39	39	39	39	39	39
N	12,169	12,169	12,169	12,169	12,169	12,169

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

Table A.4: Heterogenous effects

	(1) Big City	(2) Poor	(3) Rainy	(4) Big Subway	(5) Asia	(6) High AOD
1-18 months post	-0.0231**	-0.0227**	-0.0204*	-0.0204**	-0.0145*	-0.0124
	(0.0099)	(0.0095)	(0.0102)	(0.0095)	(0.0078)	(0.0077)
Interaction	0.0046	0.0036	-0.0010	-0.0010	-0.0108	-0.0170
	(0.0144)	(0.0137)	(0.0138)	(0.0139)	(0.0116)	(0.0112)
Mean dep. var.	0.457	0.457	0.457	0.457	0.457	0.457
R-squared	0.834	0.834	0.834	0.834	0.834	0.834
Number of events	39	39	39	39	39	39
N	12,169	12,169	12,169	12,169	12,169	12,169

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators, city population and country GDP. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

Table A.5: Placebo city AOD for 18 month period post system opening

	(1)	(2)	(3)	(4)	(5)	(6)
1-18 months post	0.0048	0.0066	0.0040	0.0058	-0.0039	-0.0014
	(0.0103)	(0.0112)	(0.0095)	(0.0104)	(0.0088)	(0.0092)
city pop.& GDP	No	Yes	No	Yes	No	Yes
city-level trends	No	No	Yes	Yes	No	No
city-level pre/post trends	No	No	No	No	Yes	Yes
Mean dep. var.	0.436	0.440	0.436	0.440	0.436	0.440
R-squared	0.817	0.819	0.822	0.822	0.823	0.824
Number of events	27	25	27	25	27	25
N	8,105	7,528	8,105	7,528	8,105	7,528

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

Table A.6: Robustness check using an expanded sample of cities and country by month fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
1-18 months post	-0.0183*	-0.0190*	-0.0137	-0.0154*	-0.0185**	-0.0202**
	(0.00953)	(0.00985)	(0.00866)	(0.00914)	(0.00799)	(0.00856)
city pop.& GDP	No	Yes	No	Yes	No	Yes
city-level trends	No	No	Yes	Yes	No	No
city-level pre/post trends	No	No	No	No	Yes	Yes
Mean dep. var.	0.431	0.433	0.431	0.433	0.431	0.433
R-squared	0.836	0.835	0.839	0.838	0.839	0.838
Number of events	39	39	39	39	39	39
N	198,381	193,867	198,381	193,867	198,381	193,867

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, country×month indicators, calendar month×city indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

Table A.7: Robustness check excluding observations with low pixel count

	(1)	(2)	(3)	(4)	(5)	(6)
1-18 months post	-0.0199**	-0.0205**	-0.0196**	-0.0208**	-0.0173*	-0.0185*
	(0.00964)	(0.00979)	(0.00922)	(0.00966)	(0.00963)	(0.0102)
city pop.& GDP	No	Yes	No	Yes	No	Yes
city-level trends	No	No	Yes	Yes	No	No
city-level pre/post trends	No	No	No	No	Yes	Yes
Mean dep. var.	0.442	0.442	0.442	0.442	0.442	0.442
R-squared	0.866	0.867	0.871	0.871	0.872	0.873
Number of events	39	39	39	39	39	39
N	10,957	10,957	10,957	10,957	10,957	10,957

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center for a constant sample of cities. Standard errors clustered at the city level in parentheses. The sample excludes observations with a number of pixel-days used to calculate AOD that falls below the tenth percentile of the distribution for a given satellite. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month×city indicators, year×continent indicators, pre- and post-period indicators. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

Table A.8: Expansions

	Main Cities		All Cities	
	(1)	(2)	(3)	(4)
Expansions	2nd	$\geq 2^{nd}$	2nd	$\geq 2^{nd}$
1-18 months post	-0.0147	0.0008	-0.0113	-0.0016
	(0.0141)	(0.0085)	(0.0086)	(0.0033)
Mean dep. var.	0.613	0.633	0.516	0.488
Mean expansion size (stations)	17.7	15.3	14.4	11.0
R-squared	0.778	0.807	0.816	0.816
Number of events	14	21	26	104
N	4,375	6,678	8,110	31,323

*Note: Dependent variable is monthly mean AOD in a 10km disk centered around the city center. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), AOD pixel count, satellite indicator, calendar month × city indicators, year × continent indicators, pre- and post-period indicators, city population and country GDP. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*

Table A.9: Results on ridership per capita

	(1)	(2)	(3)
Event	1st	2nd	$\geq 2^{nd}$
1-18 months post	62.42*	44.62***	29.03*
	(34.19)	(8.092)	(13.80)
Mean dep. var.	32.63	37.14	37.50
R-squared	0.488	0.942	0.937
Number of events	28	8	15
N	4,788	1,282	2,112

*Note: Dependent variable is ridership per 1000 of population. Standard errors clustered at the city level in parentheses. All regressions include the following controls; linear and quadratic climate controls (temperature, vapor, cloud cover, precipitation, frost days), calendar month × city indicators, year × continent indicators, pre- and post-period indicators, city population and country GDP. Stars denote significance levels: * 0.10, ** 0.05, *** 0.01.*