

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, Revised November 2020

We are grateful to Christopher Balette, Jessica Burley, Alejandra Hernandez, Tasnia Hussain, Andy Liu, Fern Ramoutar, Mahdy Saddradini, Mohamed Salat, David Valdes, Farhan Yahya, and Guan Yi for research assistance. 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.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Nicolas Gendron-Carrier, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A. Turner. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

## Subways and Urban Air Pollution

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

NBER Working Paper No. 24183

January 2018, Revised November 2020

JEL No. L91,R11,R14,R4

### **ABSTRACT**

We investigate the effect of subway system openings on urban air pollution. On average, particulate concentrations are unchanged by subway openings. For cities with higher initial pollution levels, subway openings reduce particulates by 4% in the area surrounding a city center. The effect decays with distance to city center and persists over the longest time horizon that we can measure with our data, about four years. For highly polluted cities, we estimate that a new subway system provides an external mortality benefit of about \$1b per year. For less polluted cities, the effect is indistinguishable from zero. Back of the envelope cost estimates suggest that reduced mortality due to lower air pollution offsets a substantial share of the construction costs of subways.

Nicolas Gendron-Carrier  
Department of Economics  
McGill University  
855 Sherbrooke St. West  
Montreal, QC H3A 2T7  
Canada  
nicolas.gendron-carrier@mcgill.ca

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

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

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

## 1. Introduction

Traffic congestion is a serious issue in many cities around the world, in part because of its impact on air pollution and human health. One proposed solution is subways and other public transit infrastructure. These are costly up-front investments, and there is little evidence about whether they are worth it. In this paper, we investigate the effect of subway 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 February 2000 and December 2017. 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. We consider multiple geographies, but our preferred outcome variable is air pollution in a 10km disk with centroid in the city center.

For an average opening, the data indicate that subways result in a tiny decrease in AOD during the 18 post opening months but this decrease cannot be distinguished from zero. This average effect, however, conceals considerable heterogeneity across cities. Our main analysis focuses on the 58 subway openings that occurred between August 2001 and July of 2016. Of these, our event study research design indicates that 20 openings increased AOD and 12 had no measurable effect. In the remaining 26 cities, not quite half the sample, AOD fell after the subway opened.

Our investigations fail to find measurable relationships between estimated subway effects and initial city or subway characteristics, save one. The 23 cities where subways reduced AOD are overwhelmingly among those whose initial level of AOD was above median. For the set of cities in the top half of the initial AOD distribution, with above 0.36 AOD (approximately  $28 \mu\text{g}/\text{m}^3$  PM<sub>2.5</sub>) on average in 2000, AOD levels fell by about 4% after the opening of the subway. This effect is robust to a variety regression specifications and to changes in the details of how we calculate city average AOD levels. The effect is larger near the center of cities where subway service is usually concentrated and we cease to find significant pollution effects beyond 25km from the city center. Data limitations hinder our ability to estimate the effects of subways over longer time horizons. However, the available data suggest the 4% decline in AOD experienced by an average high AOD city persists for at least four years.

As a complementary analysis, we provide evidence that subway ridership is a key correlate of AOD reductions, that is, cities where AOD fell significantly are also those in which ridership - and hence automobile substitution - is largest. Finally, we also investigate the effect of subway expansions<sup>1</sup> beyond the inaugural lines and find small effects, both for the average city and high

---

<sup>1</sup>We define expansions as the opening of new subway lines beyond the inaugural lines. This definition excludes a common expansion type in which a one or more subway stations are *added* to an existing subway line.

AOD cities.

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 58 subway openings and 143 expansions that occurred anywhere in the world between August 2001 and July 2016. Thus, we dramatically improve on our ability to assess whether subway openings, in fact, reduce 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 initially in the top half of the AOD distribution, a subway opening prevents 22.5 infant and 500 total deaths per year. Using standard income-adjusted life values, this averted mortality is worth about \$43m and \$1b 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. Note that these estimates apply only to cities that are initially in the upper half of our AOD distribution. For cities in the lower half of the AOD distribution, the effect of subway openings is small and not distinguishable from zero.

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 PM<sub>10</sub>, 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. 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 (2018) organized into a monthly panel and updated to December 2017. These data define a ‘subway’ as an electric-powered urban rail system isolated from interactions with automobile traffic and pedestrians. This excludes most streetcars because they interact with vehicle and pedestrian traffic at stoplights and crossings, but 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 originally compiled these data manually between January 2012 and February 2014 and updated them during 2020, using the following process. First, using online sources such as <http://www.urbanrail.net/> and links therein, together with links on wikipedia, we compiled a list of all subway stations worldwide. Next, for each station on our list, we recorded 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 2017, the time period for which our air pollution data is available.

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 2017, to base this exercise on a constant sample of cities, we restrict our attention to subways that open between August 2001 and July 2016. Table A.7 lists the 58 subway system openings taking place in this time period, by date of opening, together with basic information

about the cities where they are located. Subways opened fairly uniformly throughout the period and the average opening date is June 2010.<sup>2</sup>

For 42 of these 58 cities we are able to gather ridership data describing unlinked trips, mostly from annual reports, statistical agencies or industry associations.<sup>3</sup> Ridership is reported at the monthly level for around half of these 42 cities. For the rest, we interpolate to calculate monthly ridership from quarterly or yearly data. Table A.7 also reports mean daily ridership for each city where data is available, 18 months after the system’s inauguration. For the average city in our sample, 130,000 people rode the subway on an average day in the eighteenth 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 about triples over the first three years of operation, from about 20 to 60 riders per thousand of population. The rate of increase begins to slow about three years after opening.

Finally, we determine the date when construction began for each subway opening in our sample. On average, construction begins 77 months prior to opening and shows high variance.

## B Satellite-based pollution

The Moderate Resolution Imaging Spectroradiometers aboard the Terra and Aqua earth-observing satellites provide daily measures of the aerosol optical depth of the atmosphere at a 3km spatial resolution everywhere in the world (Levy, Hsu, and al., 2015a,b). Remer, Levy, and Munchak (2013) 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.<sup>4</sup> 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, 2017 for the Terra satellite, and from July 4, 2002 until December 31, 2017 for Aqua. During September of 2018, 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.

<sup>2</sup>There are 11 cities whose subway system opening occurs within 18 months of the beginning or end of the available pollution data which are excluded from the table since they are excluded from our main estimation sample.

<sup>3</sup>Table A.1 reports data sources for ridership data.

<sup>4</sup>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. The MODIS data are available for download at <ftp://1adsweb.nascom.nasa.gov/allData/6/>.

Figure 2 presents a map showing AOD for December 1, 2017 and for average AOD over 2017, both from the Terra satellite. Darker shades indicate 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 December 1, 2017 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 is less dramatic.

Figure 3 panel (a) illustrates the AOD data for Palma, Spain in June of 2017. To show scale, the large circle in this image is 25km in radius. Panel (b) provides the corresponding image for December 2017. The small white circles represent subway stations at the end of the sample.

While it is straightforward to calculate average AOD for any given area, we must choose the area over which to average. Panel (c) illustrates the geographies that we consider as the basis of this study. They are the 10km and 25km radius rings illustrated in panel (c), and the intersection of these rings with a city footprint formed by tracing around the area that is sufficiently brightly lit at night to be considered part of the urban area. Panel (c) illustrates this footprint as the gray polygons.

In Appendix C we check that AOD is highly predictive of ground-measured PM10 and PM2.5. In addition, we find that the relationship between ground-based and remotely-sensed particulates is not sensitive to the exact region we use to calculate city average AOD. That is, AOD is about equally good at explaining ground based measures if we evaluate AOD over a 10km disk centered on the city, a 25km disk centered on the city, or the corresponding city footprints derived from lights at night data. Given this, our analysis relies primarily on AOD averages calculated over 10km disks. Finally, we note that in our preferred specification, one unit of remotely sensed AOD is associated with about  $114 \mu g/m^3$  of PM10 measured by a ground-based instrument.

With daily images in hand, it is straightforward to construct monthly averages. To calculate our city level monthly AOD measure, we average over the 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  $r \in \{10km, 25km, 50km\}$  as well as the intersection of these disks with city footprints for all cities in our sample.

Table 1 provides worldwide and continental summary statistics. In 2017, the average AOD

reading within 10km of a city center from the Aqua satellite was 0.40. It was higher in Asian cities, 0.50, 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 1 also reports mean AOD over the more carefully drawn city footprints illustrated in panel (c) of Figure 3. Interestingly, these values barely differ from the corresponding values calculated on the basis of the comparable disk. Table A.7 reports the mean and standard deviation of AOD for each of our sample subway cities using the Terra satellite.

In an average month, the AOD reading for an average 10km city disk is based on 89 pixel-days for Aqua and 104 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 4% of city-months contain zero pixel-day observations and do not appear in our sample.

Table 1 also reports pixel counts for the two footprint measures. The count of pixels in the 10km disk and 10km footprint means are quite close. This is consistent with what we see in Figure 3. The lights based city footprint typically almost completely includes the 10k disk centered on the city. In contrast, the 25km footprint measure contains many fewer pixels than does the 25km disk. This is also consistent with what we see in Figure 3.

The second panel of Table 1 presents AOD averages for 2000 for 10km 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 AOD, and no obvious trend for the other continents or for the 58 cities as a whole.

We observe a strong seasonal pattern in AOD. This reflects seasonal variation in cloud cover and motivates our use of city-by-calendar month indicators as controls in our regressions. Together the two panels of Figure A.1 suggest that a relationship exists between the extent of AOD data availability 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 experiment broadly with sampling rules that reduce the importance of city-months with sparse AOD data and with controlling for the count of pixel-days used to construct each city-month average and find no qualitative changes in our results. 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***

Consistent with the large related literature that we describe in Appendix C, we find that local weather conditions are important determinants of remotely sensed 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. Together with AOD pixel count 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 city population data comes from the United Nation’s World Urbanization Prospects data (DESA Population Division, 2018). These data describe annual population counts for the universe urban agglomerations with populations exceeding 300,000 in 2018. We use these data in three main ways. First, they 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. Second, we rely on these data as source of city level population data. Third, in some of our regressions, we expand our sample to include non-subway cities as a way to increase the precision with which we estimate control variables effects. In these regressions, we include all of the 461 non-subway cities in our sample.

We use the Penn World Tables to obtain annual measures of country GDP (Feenstra, Inklaar, and Timmer, 2015) for all cities in our sample. Table A.7 reports population and country level GDP per capita for each city in our estimating sample at the time of subway system opening. Cities that opened subways in recent years are large, their average population is 4.0m when their subways open, and they tend to be in middle or high income countries.

### **3. Conceptual framework**

The logic of a relationship between subways and air pollution rests on two hypotheses about the relationships between subways, travel behavior, and pollution. First, that internal combustion engines contribute to air pollution. Second, that subway trips reduce motor vehicle trips, at least at certain times and places in a city. We discuss both below, along with other factors that have the potential to amplify or attenuate these relationships in different urban settings.

The relationship between automobile traffic and pollution is much studied. Chen, Yang, Qin, and Xu (2016) uses variation in Beijing traffic caused by a license plate restriction and widespread superstitious avoidance of licence plates ending in ‘4’ to provide convincing causal estimates of a negative relationship between a traffic congestion index and an air quality index. 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% PM10 reduction.<sup>5</sup> The variation in these point estimates foreshadows the variance we will find in the effects of subways on AOD, and it is natural to suspect that physical geography may be partly responsible for this variation. Unlike the studies described above, our sample is large enough to allow us to explore whether wind patterns and city topography are relevant determinants of heterogeneous effects of subway systems, by allowing pollutants to be easily dispersed or entrapped, respectively.

The literature also investigates the relationship between subways and vehicle traffic in large cities. Anderson (2014) examines changes in Los Angeles traffic congestion during a 35 day transit strike in 2003. During the strike, delay per mile increases by about half during peak hours on highways serving the same region as LA's subway lines. The effect on delay at other times and places is approximately zero. In a sample of 42 Chinese cities, Gu, Jiang, Zhang, and Zou (2020) examine changes in a cell-phone-based measure of street-segment-by-hour traffic speed for segments near newly opened subway lines. They find a measurable increase in traffic speed for 48 weeks (the extent of their data) following the opening of a subway line. Both results are consistent with subway openings causing a reduction in driving.

Together, these results suggest subways affect pollution by reducing traffic. Chen and Whalley (2012) tests this relationship directly by comparing air pollution in central Taipei during the year before and after the opening of the Taipei subway in March of 1996. Comparing hourly pollution measurements from measuring stations in central Taipei to hourly ridership data, they estimate a 5%-15% reduction in Carbon Monoxide following the subway opening, about the same effect on Nitrous Oxides, but little effect on either Ozone or particulates.

This evidence in hand, we note its limits. The results above describe time periods measured in weeks and are restricted to routes that compete with the subway. Given the generational planning horizon for subway systems, extending these results to a longer time horizon is important. Over a longer time horizon, other responses to subway availability may offset short run reductions in traffic and pollution. Foremost is the possibility that decreases in traffic caused by the subway will be short-lived as drivers eventually exploit the newly available road capacity. Duranton and Turner (2011) provide indirect evidence for such an effect. Comparing total vehicle miles driven in US metropolitan areas to total lane miles of highways over a ten year period, they find that a 1% increase in highways causes a 1% increase in total driving.<sup>6</sup> Although Duranton and Turner (2011) studies highway expansions, the logic of their finding suggests that a reduction in traffic due to

---

<sup>5</sup>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.

<sup>6</sup>Hsu and Zhang (2014) replicate the Duranton and Turner (2011) for Japan, Garcia-Lopez, Pasidis, and Viladecans-Marsal (2017) for Europe, and Chen and Klaiber (2020) for China.

increased subway ridership, like other capacity expansions, should cause an offsetting increase in the demand for automobile and truck travel. If so, subways will not reduce traffic and pollution in metropolitan areas over a decennial time scale. We will show well-identified short-term effects as well as evidence from longer time horizons to shed some light on the issue of temporal decay.

To the extent that subways reduce the cost of urban mobility we should also expect changes in the spatial organization of economic activity. This is the most basic prediction of theoretical models of cities (Duranton and Puga, 2015) and the evidence that transportation infrastructure in general, and subways in particular, lead to a reorganization of cities is compelling. Baum-Snow (2007) finds that the interstate highway system played an important role in the decentralization of US cities from 1950 to 1990. Baum-Snow, Brandt, Henderson, Turner, and Zhang (2017) confirm that Chinese highways have much the same effect on Chinese cities. For subways, in a sample of 161 subway cities between 1950 and 2010, Gonzalez-Navarro and Turner (2018) find that subway networks tend to favor movement in and out of the downtown core and that they cause cities to become more decentralized. Heblich, Redding, and Sturm (2020) refines this conclusion and shows that the opening of the London underground led to a concentration of employment in the center and increased separation of workplace and residence.

In short, subways reduce the cost to move people. In response, cities reorganize in ways that require greater mobility: they spread out and the separation of workplace and residence increases. In the analysis that follows, we will investigate the spatial extent of the effect of subways on air pollution. In particular, we will be interested in estimating pollution impacts at different distances from the downtown of a city. Because subways tend to serve central cities, substitution of automobile trips should confer larger pollution reductions in areas closer to downtowns and decay as distance from the downtown increases.

Finally, a central prediction of most theoretical models is that reductions in transportation costs will draw more people and economic activity into a city (e.g. Duranton and Puga (2015)). Therefore, to the extent that pollution and economic activity are complementary, subways should increase pollution levels.<sup>7</sup>

Summing up, the available evidence supports the idea that subways reduce traffic and hence pollution during (about) the year after the subway opens. This evidence, however, is pieced together from studies of small numbers of cities and it is natural to wonder whether this chain of causation leads to the same outcome for all new subways.

Back-of-the-envelope calculations of the sensitivity of automobile travel to subway openings

---

<sup>7</sup>We note, however, Gonzalez-Navarro and Turner (2018) provide contrary evidence. They examine the relationship between subway system extent and population in a sample of 161 subway cities over decennial time scales and find no relationship.

suggest that subways can reasonably provide between about 1.5% and 10% of trips in a city.<sup>8</sup> 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, a subway reduces PM10 by between 1.5% and 10%. This brackets our estimated 4% reduction in high AOD cities.

This example highlights two points. First, the logic of induced demand will mean that each 1% of trips diverted to the subway will elicit additional automobile trips. Moreover, a 1% reduction in automobile trips will generally result in a less than 1% reduction in pollution. Therefore, given the back of the envelop calculations described above, if subway openings result in meaningful reduction on AOD, it will be because they are displacing trips that are particularly dirty, either because they involve particularly dirty vehicles or trips that make large contributions to congestion.

Second, the range of effects suggested by our back of envelope calculations suggests that the way cities respond to subways is likely to be heterogeneous. Indeed, in our sample, ridership varies from a low of 4,000 passengers per day to a maximum of over 400,000 passengers per day 18 months after opening. Understanding this heterogeneity will be an important part of our analysis. If subways are to play a role in improving urban air quality, it is important to learn where we should expect larger effects. On the basis of the discussion above, our investigation of city heterogeneity implicitly involves looking for cities where subways draw a lot of riders, where induced demand is relatively small, and where car trips are relatively polluting.

More specifically, we investigate how the effect of subways on AOD varies with the following classes of city attributes. Cities that already enjoy clean air have less potential to improve, while windy cities, and those with flat geography should be better able to disperse pollution. Poor countries may benefit more from subway construction given that their vehicle fleets tend to be more polluting. City size may explain heterogeneous effects because traffic congestion is most relevant in cities experiencing large or increasing levels of population. Thus, we investigate heterogeneous effects by city population density, city size, city growth and whether the subway system at opening is relatively large or small.

---

<sup>8</sup>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.7, in the year of subway opening an average city in our sample has a population of about 4.0m. 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 50 rides per 1000 of population. Applying this rate to a city of 4.0m, 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. An average US household took 3.8 trips per day, (2009 National Household Transportation Survey, [nhts.ornl.gov/2009/pub/stt.pdf](https://nhts.ornl.gov/2009/pub/stt.pdf), Table3.) about 90% by car (Duranton and Turner, 2018). If everyone in a city of 4.0m 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 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%.

## 4. Results

### A Econometric framework

In our primary econometric exercise, we examine changes in AOD using an event study design that exploits the high frequency air pollution data as well as the sharp starting date of subway systems.

Let  $i = 1, \dots, I$  index cities and  $t$  index months. Let  $AOD_{it}$  denote AOD in city  $i$  at time  $t$ . We usually observe a 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,  $\tau_{it}$  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.,  $\tau_{it} \in \{-k, \dots, 0, \dots, k\}$ . 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. Now define the following families of indicator variables,

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

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

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.

To estimate the effect of subway system opening on urban air pollution we use the following event study specification:

$$AOD_{it} = \beta_i + \sum_{j \in \{-k, \dots, k\}} \alpha_j D_{it}(j) + \gamma' X_{it} + \epsilon_{it}, \quad (3)$$

where  $\beta_i$  is a city fixed effect and  $D_{it}(j)$  is a dummy indicating observations  $j$  months from a subway opening in city  $i$ . The reference category excluded from the regression is  $D_{it}(-6)$ .  $\alpha_j$  are the coefficients of interest, which will be negative after opening ( $j > 0$ ) if subways improve air quality. We are also interested in verifying that the estimated coefficients corresponding to months before subway opening ( $j < 0$ ) are not statistically different from zero as this provides evidence of a lack of pre-existing trends in pollution before treatment. We can also bin the data in a more aggregated manner so that the coefficient  $\alpha_j$  can correspond to a series of three or six month bins in Equation (3). When we do this, the reference period becomes the bin containing month 6 before a subway opens.

We also include the following control variables: A satellite indicator to account for slightly different AOD measurements from each satellite, year-by-continent indicators to flexibly account for regional trends in AOD, and city-by-calendar month (1-12) indicators to capture seasonality

in pollution patterns that repeat annually but are idiosyncratic to each city. We control for linear and quadratic terms in monthly city-specific climate variables (cloud cover percentage, frost day frequency, mean temperature, precipitation, and vapour pressure). Finally, we also control for the count of AOD pixel-days used to calculate the city-month AOD measure.

For most of our analysis, we use a 37 month study window, running from 18 months before, to 18 months after each city’s subway opening. 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. Here and throughout, we cluster errors at the city level.

We also allow for a city specific pre- and post- treatment window by including city specific dummies  $\mathbb{1}_{it}(\tau < -k)$ ,  $\mathbb{1}_{it}(\tau > k)$  and  $D_{it}(0)$  as part of our control set. This allows us to 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 estimate the effect of subways on AOD. Month of opening  $\tau = 0$  is ‘partially treated’ depending on the exact day subway system started making it harder to interpret.<sup>9</sup>

In addition to the event study graph, we also present regression results in table format. Here, instead of estimating a series of coefficients around subways introduction, we compare AOD in the 18-month period before the subway to the 18-month period after it starts. More specifically, these regressions take the form:

$$AOD_{it} = \beta_i + \alpha_1 D_{it}(1, k) + \gamma' X_{it} + \epsilon_{it}. \quad (4)$$

This is a regression of AOD on an indicator for post-opening and controls, keeping everything else unchanged from the event study specification. The coefficient of interest is  $\alpha_1$ , the difference in conditional mean AOD between city-months with  $\tau_{it} \in \{-18, \dots, -1\}$  and those with  $\tau_{it} \in \{1, \dots, 18\}$ .

## B Average Effect

We present our estimates of the average effect of subways on air quality in Figure 4. The figure displays estimated coefficients  $\alpha_j$  around subway opening from Equation (3) as well as 95% confidence intervals. The vertical axis is AOD units. Panel (a) shows three month bin estimates while panel (b) presents results using six month bins, where the indexing of  $\alpha_j$  runs over these sets of bins. As explained above, the bin containing month -6 is set as the reference period and the point estimate is depicted in the middle of the bin. For example, in the 3 month bin specification shown in panel (a) the reference bin refers to months -4, -5 and -6 and the point estimate is shown in the middle month of the bin ( $\tau=-5$ ). Both figures display a lack of pre-treatment trends as well as a

---

<sup>9</sup>In an earlier version of this paper, we used common pre- and post-indicators. These coefficients reflect sample mean pre- and post period AOD averages. This is problematic conceptually because these variables could, in theory, also measure treatment effects. It is problematic in practice because these pre- and post- period AOD averages sometimes change across samples and (slightly) affect estimates of treatment effects. The more flexible specification resolves this problem.

slight reduction in AOD that is not statistically different from zero. This figure suggests that for the average city in the sample we cannot reject that the effect of subways on urban air pollution is nil. Note that mean AOD in our sample is about 0.46, so monthly variation in point estimates is small relative to the mean level.

Table 2 shows results using specification (4). Moving across the columns of the table we add progressively more controls. Column 1 estimates a version of regression 4 with a minimal set of controls.<sup>10</sup> Column 2 adds satellite and continent by year dummies, Column 3 adds city by calendar month dummies, Column 4 adds climate controls, while Column 5 instead uses climate by continent controls. In all cases we estimate a treatment effect of about -0.01. This is small relative to the 0.46 sample mean AOD and is not distinguishable from zero at conventional levels of significance. Column 6 adds observations from all cities in the UN cities database to estimate more precisely common controls such as continent by year and climate and we find virtually the same result.

Column 7 investigates whether we observe a change in the time trend of AOD with subways using the main estimating sample of 58 cities and allowing for city specific time trends in AOD. The regressor of interest is now an 18 month post subway dummy as well as the interaction with time. Again, while the point estimate on the AOD trend post subway opening is negative, the estimate is not sufficiently large to differentiate it from zero.

First, note that to the extent that subways decrease pollution for an average system opening in this sample, this effect is small. Second, the controls matter in the way we would hope. As we move from Column (1) to Column (5), the coefficient of interest is steady while  $R^2$ s increase and standard errors decrease or stay constant. Third, when we control for trends in Column (7), the level effect of subways shrinks and becomes slightly positive. That the trend variable is slightly negative suggests that the slight negative effect we see in the first five columns is captured by the trend variables.

Appendix Table A.3 reports robustness checks for the regression reported in Column 5 of Table 2. In particular, we repeat this specification but: (1) weight observations by the count of observed AOD pixels that make up each city-month AOD measurement count; (2) drop observations with low pixel count; (3) calculate city month AOD on the basis of three different city footprints. Appendix Table A.4 repeats the specification of Table 2 Column 5 for a variety different windows of analysis ( $k = 6$  months, 12 months, 24 months and 36 months). All of these alternative specifications confirm the conclusion suggested by Table 2: the average effect of subways on AOD is consistently negative and not distinguishable from zero at conventional levels of significance.

Following Andrews (1993), Hansen (2000), and Andrews (2003), we checked for a structural break in the level and trend of AOD around the time of system opening by estimating a series of regressions of AOD on a step function or spline as we allow the timing of the step or spline to

---

<sup>10</sup>City fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-o indicators

traverse the study period. The details of these tests are presented in the supplemental Appendix F. These tests do not support the hypothesis that there is a break in the level or trend of AOD around the time of an average subway opening.

We also note that in theory, our research design could confound the effects of system openings with those of expansions that occur soon afterward. Among the 58 cities that make up our main sample, 19 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.

### *C Heterogeneous effects*

We estimate a version of Equation (4) where we allow the effect of subways to be different in every city. When we do this, we find evidence of substantial heterogeneity of impacts. There are 20 cities where subway openings appear to cause statistically significant increases in AOD, 26 where they appear to cause a decrease, and 12 where the effect is, like the average, indistinguishable from zero.

We now turn to an investigation of the reasons for such dramatically heterogeneous responses. Table 3 investigates whether subways have different effects on different types of cities. In this table, we replicate the results of Column 5 of Table 2, but add an interaction between the post-subway indicator and a particular city characteristic.

We choose interaction effects to investigate on the basis of the discussion in Section 3. We are interested in whether and how the effect of subways on AOD varies with the overall level of hydrocarbon powered travel or the characteristics of the vehicle fleet, with the extent to which such traffic is congested, with the extent or importance of the subway, and with physical geography related to pollution dispersion.

In the first three columns of Table 3 we rely on indicator variables describing basic city characteristics; (1) above median city population in 2000; (2) above median density in 2000<sup>11</sup> and (3) above median population growth, 1990-2000. Columns (4) and (5) are country characteristics; Column (4) explores heterogeneity by median country GDP per capita in 2000; and, Column (5) whether the city is in China. China may be different in that it builds large systems very quickly in large and fast growing cities. Columns (6) and (7) investigate the role of geographical features. Column (6) explores heterogeneity by city relief and Column (7) by wind speed.<sup>12</sup> Column (8) investigates heterogeneity by the size of the subway supply shock as measured by an indicator of whether the number of stations per capita is above the median in the sample.

Column (1) suggests that subways have a larger effect on AOD in smaller cities, while Column (3) suggests that subways may have a larger effect in faster growing cities. Smaller cities tend to

---

<sup>11</sup>The density measure is the elasticity of lights at night to distance from the city center and is taken from Gonzalez-Navarro and Turner (2018).

<sup>12</sup>City relief is from Gonzalez-Navarro and Turner (2018) and wind speed is from Wentz, Scott, Hoffman, Leidner, Atlas, and Ardizzone (2015).

grow faster, so these results, although statistically insignificant, seem to be consistent with each other. We do not find heterogeneous effects by city density in Column (2). We also do not find meaningful heterogeneous effects by country income or whether the city is Chinese. Column (6) includes an indicator for above median elevation range in a 25km disk centered around the city center and Column (7) includes an indicator for above median wind speed in 2000. Neither effect is distinguishable from zero. Column (8) includes an indicator for above median number of stations per capita at opening. This effect, also, is not distinguishable from zero.

Finally, Column (9) considers above cities with above median AOD in 2000. This effect is large and highly significant. Subways have a much larger effect on AOD in more polluted cities. To help understand this heterogeneity in more detail, Figure 5 presents two funnel plots summarizing the 58 city-specific estimates of the effect of a subway opening on AOD. The top figure shows estimates among cities with above median AOD levels at baseline, while the bottom figure shows estimates for cities with below median AOD levels at baseline. The horizontal axis displays the coefficient estimate of the subway effect for each city. The vertical axis the corresponding standard error. Shaded regions indicate statistical significance at usual confidence levels. For example, the white area in the center indicates estimates that are not distinguishable from zero. The large extreme left and extreme right light gray-shaded areas indicate coefficient estimates that are significant at least at the 1% level. In both figures, the large triangle indicates the average effect that we estimated in Table 2 Column 5 for the overall sample. In this figure, all cities with above median AOD in 2000 are represented by solid dots while cities in the bottom half get hollow dots. The distribution of subway effects in the top figure for high AOD cities clearly lies to the left of those in the bottom figure with the less polluted cities in the sample.

Figure 6 goes further in showing the heterogeneity by initial level of AOD by presenting coefficient estimates for every city against average AOD level in 2000 on the  $x$ -axis. In the figure, dark dots are significant at the 5% level and the solid line shows a linear fit of the coefficient estimates. As can be seen, subway openings reduce air pollution more in initially more polluted cities. The dots to the right of the vertical dashed line correspond to the top graph in Figure 5. While clearly more negative on average, the figure also indicates a pattern of higher variance in the coefficient estimates among highly polluted cities.

To test the robustness of the effect of subways on pollution among high AOD cities, Table 4 replicates specifications (1)-(6) from Table 2 on the high AOD subset of cities. We see that the estimated effect of subways on AOD in this sample is about -0.03. This is the difference between the main effect and the interaction effect in Column (9) of Table 3. This magnitude is consistent across specifications and is distinguishable from zero (as opposed to the mean for less polluted cities) at low levels of significance in columns (4) and (5), where we include our most exhaustive sets of controls. In an effort to improve the precision of our estimates, in Column (6) we expand our sample to include 461 cities without subways in the final year of our sample. Since these cities do not experience subway openings they do not contribute to the estimate of the treatment effect

directly. However, they do increase the precision with which we are able to estimate continent level and climate related controls. In Column (6) of Table 4 we see that this strategy gives a more precise estimate of the subway effect, although the magnitude of the coefficient is essentially unchanged.

For the 10,896 city months in high AOD cities, AOD averages 0.66. From Table 4, the subway effect is about -0.028. Thus, our benchmark estimate for the effect of subways on AOD in polluted cities is about a 4% decline.

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. Given this, we implement the wild-cluster bootstrap procedure recommended in Cameron *et al.* (2008) and report wild cluster bootstrap p-values for the treatment effect based on 300 bootstrap replications in the lower part of the table. This turns out not to have much effect on the estimated precision of our results.

As a further test of the robustness of the effect of subways on AOD in high AOD cities, Table A.6 repeats the specifications (5) and (6) from Table 4 using treatment and control windows of 6, 12, 24 and 36 months. Effects are generally negative but are statistically distinguishable from zero only for the 36 month pre/post window and the specification of Column (6) of Table 4, that is at windows of analysis of  $\pm 18$  months or  $\pm 36$  months.

Tables 4 and A.5 indicate a robust and fairly precise estimate of the treatment effect. Table A.6 suggests more caution. With this said, we note three reasons to discount the evidence in Table A.6. First, the AOD process is clearly a noisy one, so we should not be surprised to see precision increase as we increase window width. Second, Figure 1 suggests that it may take some time before subways achieve their equilibrium usage, and so a smaller effect over short windows may actually reflect how subways work over shorter time horizons. Table 8, which uses the ridership data, will provide some support for this idea. Third, when we consider a longer treatment window, we must exclude almost a quarter of the high AOD cities. Given the cross-city heterogeneity on display in Figure 5, we should not be surprised that the average effect is sensitive to sampling.

#### **D Longer run effects in high AOD cities**

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 5 we continue to consider a pre-treatment period beginning 18 months before an opening, but consider longer post-treatment periods. Panel (a) on top reports effects by year to assess dynamics, while panel (b) below reports average effects over the post-treatment period. In columns 1 and 2 we consider two years after an opening using a specification that is otherwise the same as we in Table 2 columns 5 and 6. We see that the one year effect is about  $-0.02$ , slightly smaller, but statistically indistinguishable from our estimate of the 18 month effect in Table 2. Point estimates of the second year effect are about the same and are estimated with slightly higher precision than 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) 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  $-0.03$ . We also provide wild-cluster bootstrap p-values for the average effect. Bootstrapped p-values do not indicate a level of precision that is dramatically different from conventional estimates based on asymptotics.

In columns 3 and 4 of Table 5 we extend the post-treatment period to 36 months. Each of the three post treatment years are negative and indistinguishable from  $0.03$ . The second and third year effects are estimated with about the same precision as the one year effect but because they are larger, can be distinguished from zero. Panel (b) of Table 5 estimates the average subway effect over the three year post-period considered in columns (3) and (4). This average is about  $-0.03$  and different from zero. The bootstrapped p-value is consistent with asymptotic standard errors. In order to accommodate the longer time horizon, we need to drop 3 cities from our sample of 27 high AOD cities, so some of the difference between columns (1-2) and (3-4) may reflect this change in sample.

Finally, columns (5) and (6) extend our analysis to four years post opening. Here we see larger effects that are distinguishable from zero in three of four years. The average effect over all years is  $-0.035$ . This effect is different from zero, but not different from  $0.03$ . Again, extending our treatment period requires dropping an additional 3 high AOD cities that open their systems between three and four years of the end of our sample.

Inspection of these results reveals two patterns. First, across specifications, point estimates are within 1.64 standard errors of the 18 month effect of  $-0.03$  in almost all cases. Second, in most specifications, coefficient magnitudes increase with time from opening, while standard errors stay about constant with time from opening. These results are consistent with a persistent AOD decrease of about  $0.03$  following subway openings in high AOD cities that lasts at least four years.

### ***E Spatial scale of effect in high AOD cities***

Subways overwhelmingly serve the areas close to the most central part of a city. Gonzalez-Navarro and Turner (2018) document that about 40% of all subway stations in existing subway systems lie within 5km of the center and another 30% within 10km of the center. Thus, we expect larger effects on AOD nearer to the centers of high AOD cities than further away

Table 6 documents precisely this phenomenon. For reference, the first two columns of this table reproduce the first two columns of Table 4 which use a 10km disk around the city center to measure air pollution. Columns 3 and 4 are identical to columns 1 and 2, except that the dependent variable is mean monthly AOD in a donut between 10 and 25km of the center rather than within 10km. Columns 5 and 6 examine AOD in the donut between 25 and 50km. As expected, the effect of subways on AOD decreases with distance from the city center in our sample of high AOD cities. In particular, the estimated improvements in air quality among high AOD cities are no longer statistically different from zero beyond 25km from the center.

## F Further results

*Expansions vs openings:* Our investigation has focused on the effects of the initial opening of a subway. We now turn to an investigation of expansions, that is, the addition of subway lines after the system has entered service.

In Column (1) of Table 7 we consider all subway expansions that occur during the study period examined by Table 2. This sample includes all expansions, including those that occur in cities that open their systems prior to the beginning of our study period, and may include more than one expansion for some cities. This leaves us with a sample of 70 cities and 143 expansions. We conduct a regression like that described in Equation (4), but our ‘event’ is an expansion rather than an opening. The estimated effect of expansions is tiny ( $-0.004$ ) and indistinguishable from zero.<sup>13</sup> In the second column of Table 7 we restrict attention to all expansions that occur in one of the 58 cities that are part of our main sample. Of these 58, we observe expansions in 20. For this sample we observe a negative effect that is just distinguishable from zero. In columns (3) and (4) we repeat the first two columns, but also interact the treatment with an indicator for a high AOD city, much as in Table 3. We see that the effects of expansions are larger in high AOD cities, but this change is not distinguishable from zero.<sup>14</sup> In sum, this table provides suggestive evidence that subway expansions are on average less important than openings.

Just as for openings, our data permits us to estimate the effect of every single subway expansion on AOD. Figure A.5 presents all 70 such estimates that are possible using the sample of Column 1 of Table 7. This figure is analogous to Figure 5; coefficient estimates are on the  $x$  axis and standard errors are on the  $y$  axis. Expansion effects are statistically different from zero with progressively higher levels of significance as they lie in darker regions of the figure. This figure looks much like the corresponding figure for subway openings, although, if anything, there is even more variation

<sup>13</sup>In a handful of cases, expansions occur within 36 months of each other or of an opening. In this case, control months for one event are treatment months for another. We experimented with dropping these events and it does not qualitatively affect our results.

<sup>14</sup>To be consistent with the definition used up until Table 7, we define ‘high AOD cities’ as those with average AOD levels in 2000 at or above the median in our main sample of 58 cities with openings between August 2001 and July 2016. The median AOD level in 2000 in our sample of 58 is 0.366.

in the effects of expansions than openings. Unlike openings, however, expansions in high AOD cities are not measurably different from those in low AOD cities.

*Ridership:* Table 8 examines the effects of subway openings and expansions on subway ridership and investigates the relationship between ridership and AOD. All four regressions examine changes around the opening of the subway using regressions like Equation (4) and we restrict attention to the 42 out of 58 cities that opened a subway during our study period for which we could assemble ridership data.

In Column (1), we document average daily ridership per 1000 people in the 18 month period after system opening. In line with Figure 1, we see that an average system carries about 36 riders per thousand of population in its first year and a half of operation.

In Column (2) we replicate Column (5) of Table 2 and predict the change in AOD during the 18 months post opening. We see that, in this set of 42 cities with ridership data, subway openings decrease AOD by about 0.018 units and that the mean AOD in this sample is about as high as in the high AOD city sample. Thus, the 42 cities where ridership data are available have higher air pollution levels and, on average, have effects of subway openings on AOD that are more responsive to subways than the 17 cities where ridership data are not available.

In Column (3) we repeat Column (2) but add an interaction between the treatment indicator and an indicator for above median ridership per 1000 people. We see that the interaction effect is strongly negative. This suggests that the reduction in average AOD is driven by effects among cities where subways are used intensively. Column (4) is like (3) but we measure subway usage with an indicator for above median level of subway ridership. This estimate is very close to that of Column (3). Because ridership could be endogenous in this regression specification, some caution is required in interpreting these results. With that said, these estimates suggest that subways affect AOD by attracting riders and diverting automobile trips.

## 5. Value of AOD reductions following subway openings in high AOD cities

### 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  $\mu\text{g}/\text{m}^3$  of PM10. Thus, of 100,000 births, a 1  $\mu\text{g}/\text{m}^3$  decrease in ambient PM10 averts about 12.5 infant deaths.<sup>15</sup> 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  $\mu\text{g}/\text{m}^3$  of PM10. This estimate implies that of 100,000 births, a 1  $\mu\text{g}/\text{m}^3$  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

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

that a one unit  $\mu\text{g}/\text{m}^3$  decrease in ambient TSP averts about 5.2 infant deaths per 100,000 births. Converting from TSP to PM10 is non-trivial, however  $\text{PM10} = 0.55 \times \text{TSP}$  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  $\mu\text{g}/\text{m}^3$  decrease in ambient PM10 averts about 9.5 infant deaths per 100,000 births. In sum, these studies suggest that a one  $\mu\text{g}/\text{m}^3$  decline in PM10 averts about 10 infant deaths per 100,000 births.<sup>16</sup>

That none of these estimates can be distinguished from the others despite a range of mean PM10 of from about  $28 \mu\text{g}/\text{m}^3$  for Knittel *et al.* (2016) to about  $67 \mu\text{g}/\text{m}^3$  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  $5\text{-}100 \mu\text{g}/\text{m}^3$ , although they find non-linearities outside of this range.

Figure A.2(b) shows that most of our city-months fall in this  $5\text{-}100 \mu\text{g}/\text{m}^3$  range. Therefore, in light of the results described above, we can reasonably assume a linear dose-response relationship in our sample.

These results, together with our estimates, allow us to estimate annual infant deaths averted by a subway opening for an average high-AOD city. From Table 4, subway openings cause about a 0.028 unit decrease in AOD. Using Column 2 of Table A.2 to convert from AOD to PM10 gives about  $3.2 \mu\text{g}/\text{m}^3$  of PM10. At 10 infant deaths per 100,000 births per  $\mu\text{g}/\text{m}^3$  of PM10, the number of averted deaths due to a subway opening in city  $i$  is given by

$$3.2 \times (10 \times 10^{-5}) \times \text{Birthrate}_i \times \text{Population}_i,$$

An average city in our high AOD subset sample has a population of about 5.3 million in the year before its subway opens. With a 2% birthrate, a subway opening in this city averts about 34 infant deaths per year.<sup>17</sup>

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 the average subway opening in high AOD cities averts 22.5 infant deaths per year.

<sup>16</sup>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  $\mu\text{g}/\text{m}^3$  of PM10.

<sup>17</sup>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).

To monetize this benefit, we use country-adjusted values of a statistical life (VSL) to value averted infant deaths in each city.<sup>18</sup> Averaging over all high AOD cities, the value of averted infant deaths is \$43.0M per year. Our estimates do not allow us to conclude that subways continue to affect air quality beyond 5 years after their opening date. With a 5% discount rate, the present discounted value of this amount over five years is about 195.5M dollars. If the effect is permanent, the corresponding present value is 903M 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 D.

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 PM<sub>2.5</sub> 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 PM<sub>2.5</sub> concentrations were reduced to a counterfactual exposure level. We proceed in three steps. First, using coefficients in Column 8 of Table A.2 (Terra), we predict city-level PM<sub>2.5</sub> 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  $1.7 \mu\text{g}/\text{m}^3$  PM<sub>2.5</sub> decrease from the predicted pre-subway exposure level.<sup>19</sup> 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 high AOD cities, the 0.028 unit decrease in AOD that follows a subway opening saves about 500 lives per year.

Valuing these lives, as before, with country-adjusted values of a statistical life, the value of averted deaths is \$1B per year. With a 5% discount rate, the present discounted value of this amount over five years is about 4.54b dollars. If the effect is permanent, the corresponding present value is 21b dollars.

## **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 par-

<sup>18</sup>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.

<sup>19</sup> From Column 8 of Table A.2, a 0.028 unit decrease in AOD converts to a  $0.028 * 60.57 = 1.696 \mu\text{g}/\text{m}^3$  decrease in PM<sub>2.5</sub>.

ticulates. 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 195.5m and 903m, 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 4.54b and 21b, 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 account at least a substantial fraction of subway construction costs in cities with poor baseline air quality (above 0.36 AOD on average in a year or approximately  $28 \mu\text{g}/\text{m}^3$  PM<sub>2.5</sub>).

## 6. Conclusion

Our results indicate the that, for the 58 cities that opened a subway system between August 2001 and July of 2016, the average effect of subway system openings on AOD was slightly negative and indistinguishable from zero at ordinary levels of significance. This average conceals dramatic heterogeneity. Estimating a subway effect for each city individually, we find that in 20 cities subways appear to have caused an increase in AOD, in 12 the effect is about zero, while in the remaining 26 the effect is negative.

We find a clear pattern to this heterogeneity. Cities with initial AOD levels in the top half of our sample, that is, above 0.36 in 2000 on average, see AOD levels decrease by about 0.028 after their subways open. Since the mean AOD in this sample of cities is about 0.66, this is about a 4% decline. This effect is not sensitive to the inclusion or exclusion of a range of control variables in our regression, nor to the exact geography we use to measure each city's AOD. As might be expected, subways have a larger effect on AOD nearer the city center than further away and the effect is strongly increasing in ridership. Our ability to extend our estimates to time periods far in the future of subway opening is limited by data availability and our event study research design. With this said, subways appear to have a constant or increasing effect on AOD over the first four years of their operation. In all, the data seem to provide considerable support for the hypothesis that subway cause a modest, persistent reduction in AOD levels for cities in the top half of the initial AOD distribution, and no effect otherwise. We note that our estimates are somewhat sensitive to

the duration of treatment and control periods. We suspect that this sensitivity reflects the intrinsic noisiness of the AOD process and sampling issues that inevitably follow changes in the duration of these periods.

Our 0.028 estimated subway effect in high AOD cities is only about 4% of the level, but it is large compared to world trends. Comparing AOD levels across continents in Table 1, we see that the difference between Europe and North America in 2017 is about 0.05, one and a half subway effects, and between Europe and Asia the difference is 0.33, about 11 subway effects. In all, this suggests that subways may play a moderately important role in explaining time series and cross-sectional patterns in urban AOD.

The effects of subway expansions seem to be smaller than those of subway openings, although like openings, there is dramatic heterogeneity in these effects. Unlike openings, however, we cannot discern a pattern on the heterogeneity of the effects of expansions. In particular, subway expansions in high AOD cities are not more likely to reduce AOD than they are in low AOD cities. 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 (2018).

Extant estimates of the effects of particulates on mortality suggest that they are sufficiently poisonous that, for high AOD cities, the small nominal reductions from subway openings are economically important. In particular, they appear to be large enough to justify at least moderate subsidies for subway construction. As we should expect, our results suggest that subway openings affect AOD through their effect on ridership. This, in turn suggests that they reduce AOD by reducing the use of competing modes of transportation. This conclusion is broadly consistent with other results in the literature.

## 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* : 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. 2007. Did highways cause suburbanization? *The Quarterly Journal of Economics* 122(2): 775–805.
- Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. Roads, Railroads, and Decentralization of Chinese Cities. *The Review of Economics and Statistics* 99(3): 435–448. URL <https://ideas.repec.org/a/tpq/restat/v99y2017i3p435-448.html>.
- 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.
- 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, Shuai, Jun Yang, Ping Qin, and Jintao Xu. 2016. Truth behind chinese superstition: Non-linear effects of vehicle traffic on urban air quality in beijing. *Environment for Development Discussion Paper* (16–2).
- Chen, Wei and H. Allen Klaiber. 2020. Does road expansion induce traffic? an evaluation of vehicle-kilometers traveled in china. *Journal of Environmental Economics and Management* 104: 102387. URL <http://www.sciencedirect.com/science/article/pii/S0095069620301108>.

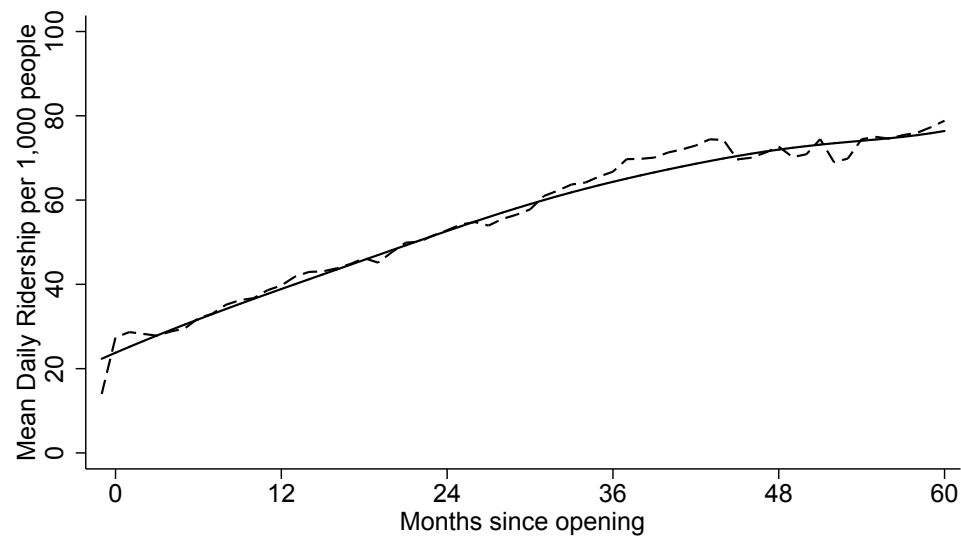
- Chen, Yihsu and Alexander Whalley. 2012. Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal* 4(1): 58–97.
- Couture, Victor, Gilles Duranton, and Matthew A. Turner. 2018. Speed. *The Review of Economics and Statistics* 100(4): 725–739. URL <https://ideas.repec.org/a/tpr/restat/v100y2018i4p725-739.html>.
- 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. 2018. World urbanization prospects: the 2018 revision.
- Duranton, Gilles and Diego Puga. 2015. Urban land use. In *Handbook of regional and urban economics*, volume 5. 467–560.
- 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. 2018. Urban form and driving: Evidence from US cities. *Journal of Urban Economics* 108(C): 170–191. URL <https://ideas.repec.org/a/eee/juecon/v108y2018icp170-191.html>.
- 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 <https://www.rug.nl/ggdc/productivity/pwt/> Accessed: 2019-09-16.
- 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*.
- 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. 2018. Subways and urban growth: Evidence from earth. *Journal of Urban Economics* 108: 85–106.
- Gu, Yizhen, Chang Jiang, Junfu Zhang, and Ben Zou. 2020. Subways and road congestion. *American Economic Journal: Applied Economics*.
- 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 ts3.26 dataset. *International Journal of Climatology*, 34(3), 623–642. Available for download at <https://crudata.uea.ac.uk/cru/data/hrg/> Accessed: 2018-10-28.

- Heblich, Stephan, Stephen J Redding, and Daniel M Sturm. 2020. The Making of the Modern Metropolis: Evidence from London\*. *The Quarterly Journal of Economics* 135(4): 2059–2133. URL <https://doi.org/10.1093/qje/qjaa014>.
- 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 wild-fires. *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 pm<sub>2.5</sub> 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, Christina Hsu, and al. 2015a. Modis atmosphere l2 aerosol product, nasa modis adaptive processing system, goddard space flight center, usa. [http://dx.doi.org/10.5067/MODIS/MOD04\\_L2.006](http://dx.doi.org/10.5067/MODIS/MOD04_L2.006). Accessed: November 2018.
- Levy, Rob, Christina Hsu, and al. 2015b. Modis atmosphere l2 aerosol product, nasa modis adaptive processing system, goddard space flight center, usa. [http://dx.doi.org/10.5067/MODIS/MYD04\\_L2.006](http://dx.doi.org/10.5067/MODIS/MYD04_L2.006). Accessed: November 2018.
- 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. Mattoo and 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.
- 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.
- Wentz, F., J. Scott, R. Hoffman, M. Leidner, R. Atlas, and J. Ardizzone. 2015. Remote sensing systems cross-calibrated multi-platform (ccmp) 6-hourly ocean vector wind analysis product on 0.25 deg grid, version 2.0. Remote Sensing Systems, Santa Rosa, CA. Available online at [www.remss.com/measurements/ccmp](http://www.remss.com/measurements/ccmp). Accessed: 2019-10-28.
- 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/health\\_topics/outdoorair/databases/cities/en/](http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/). Accessed: 2017-04-10.
- 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.

Figures

Figure 1: Daily Ridership per Capita

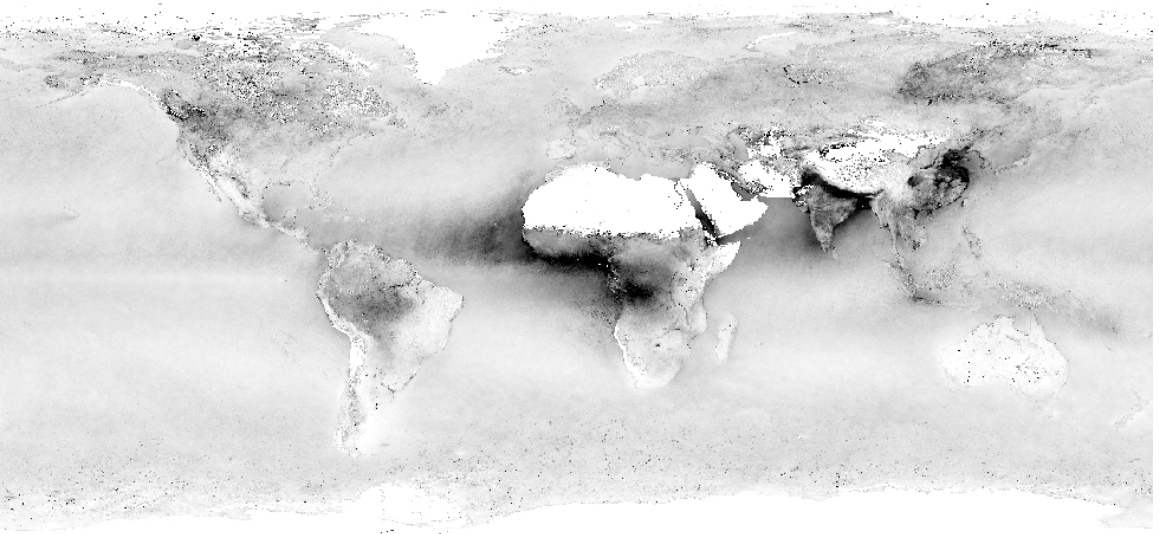


*Note: Graph depicts average daily passengers on subway per 1,000 people in metropolitan area, as well as a locally weighted regression of the series.*

Figure 2: Worldwide AOD: Terra Satellite



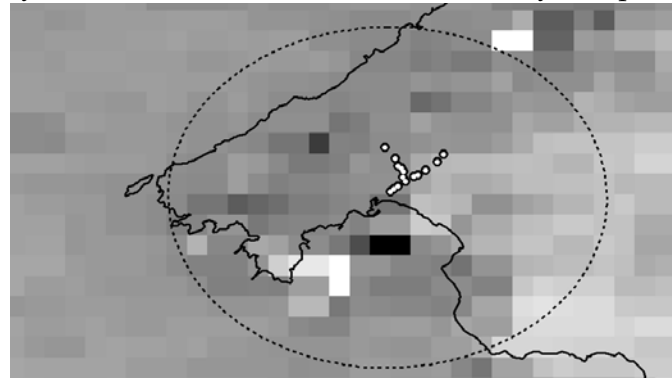
(a) Single day AOD - December 1st, 2017



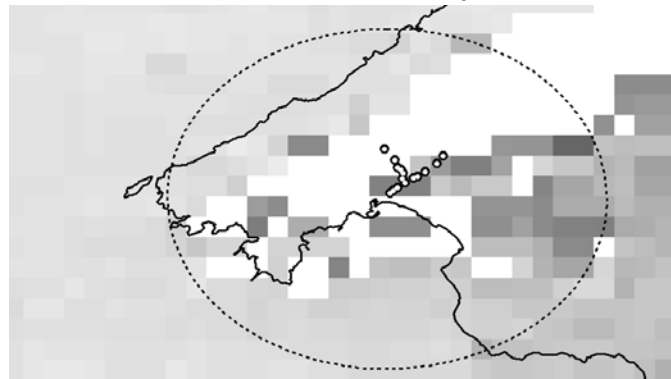
(b) Annual average AOD - 2017

*Note: Darker colors indicate higher levels of AOD. White is missing.*

Figure 3: Seasonality in AOD, Influence Area Disks, and City Footprints for Palma, Spain



(a) Palma, June 2017



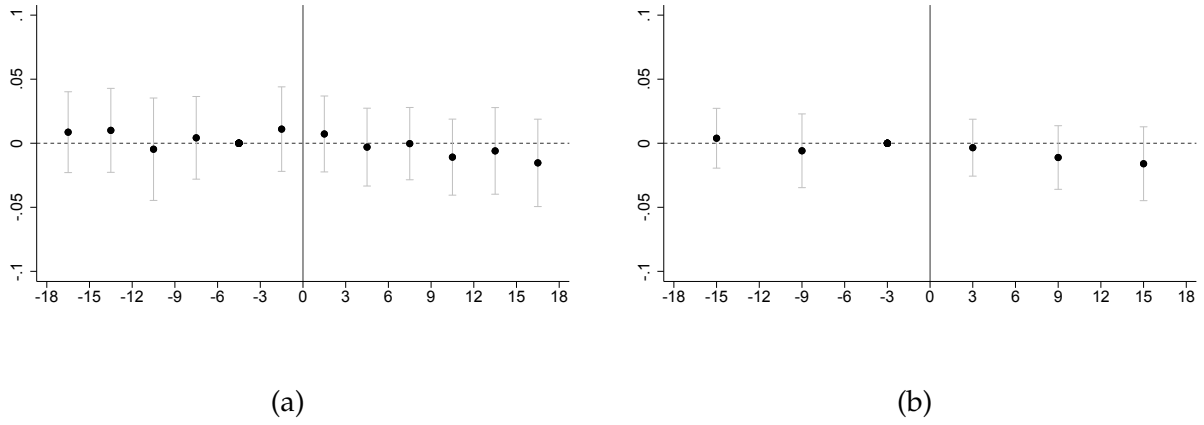
(b) Palma, December 2017



(c) Palma, extent

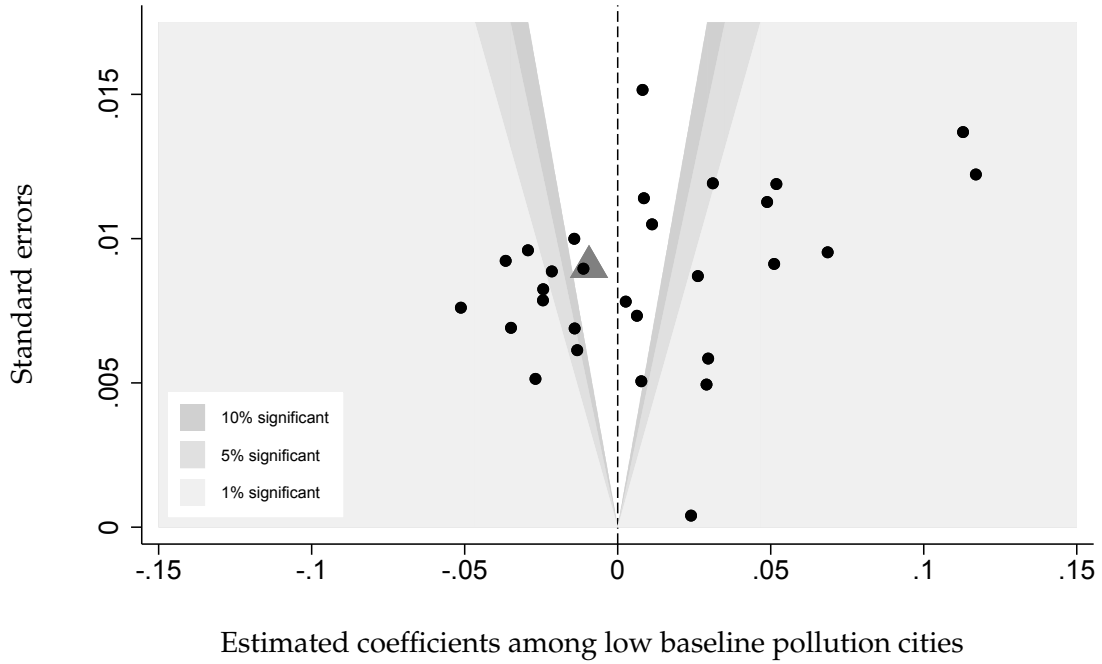
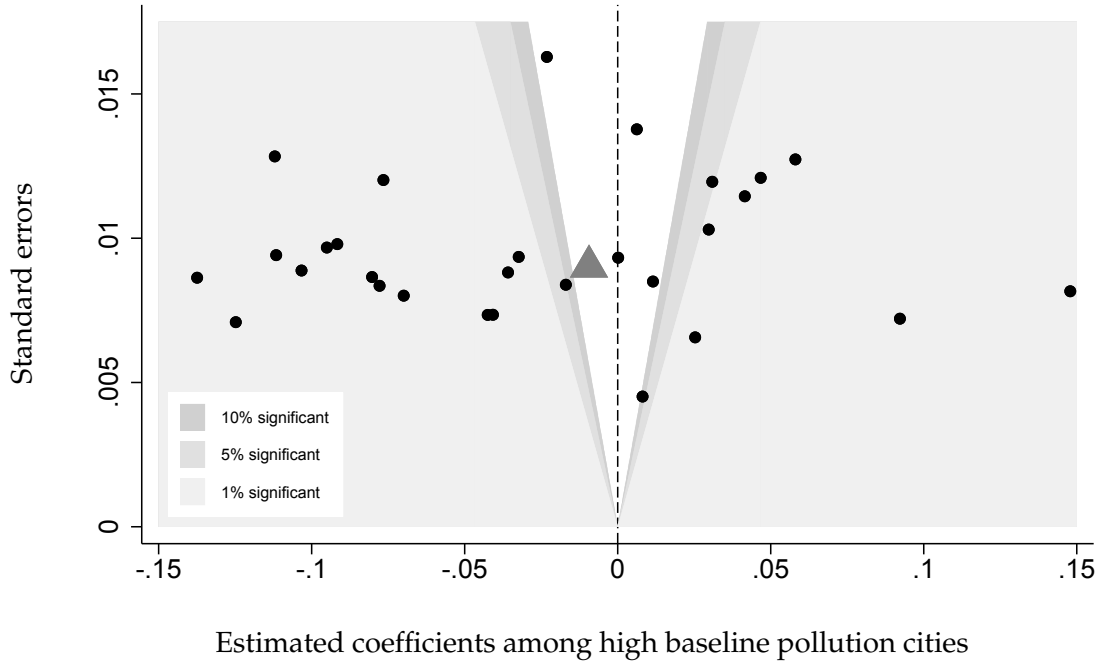
*Note: Terra AOD for Palma, Spain, in June 2017 presented in Panel (a) and December 2017 in Panel (b). Darker pixels indicate areas where AOD is higher and white pixels indicate missing values. The large circle in each image has a radius of 25km and is centered on the central business district as reported in DESA Population Division (2018). Panel (c) illustrates the 25km disk from above, the 10km disk on which our analysis usually relies, and the lights at night-based city footprint. Subway stations as of December 2017 are shown as small white circles.*

Figure 4: Event Study



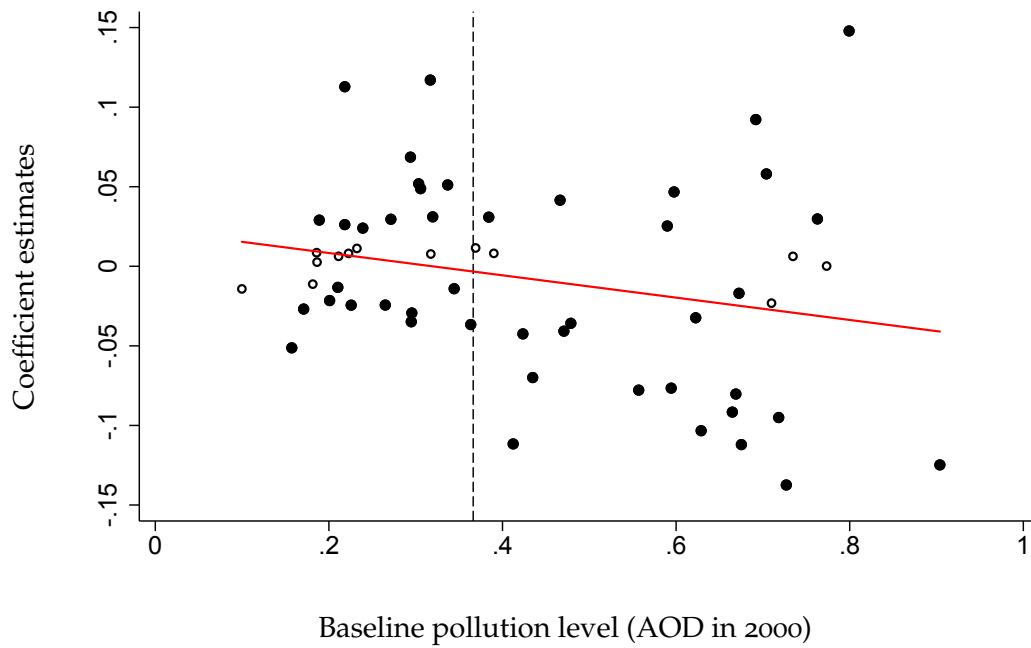
*Note: Figure shows event study coefficients of AOD around the time of subway opening and confidence intervals based on standard errors clustered at the city level. The vertical line at  $x = 0$  indicates the month of subway opening. Panel (a) shows coefficients estimated using three month bins (reference bin is months  $\tau = \{-6, -5, -4\}$ ). Panel (b) shows coefficients using six month bins (reference bin is months  $\tau = \{-6, -5, -4, -3, -2, -1\}$ ). Coefficient estimates are shown in the graph at the midpoint of the bin. Controls are as in Column 5 of Table 2.*

Figure 5: Heterogeneous Treatment Effects by High and Low Baseline AOD Levels



*Note: These funnel graphs show the 58 city-specific subway effects as small circles based on Equation (4) where the coefficient  $\alpha$  is allowed to vary by city. The x-axis shows coefficient estimates, the y-axis shows standard errors. The region in white contains estimates that are not significantly different from zero. Dark, medium, and light gray regions are different from zero at 10%, 5% and 1% in two-sided tests, respectively. The top (bottom) panel shows estimates for cities with above (below) median AOD levels at baseline. In both figures, the large triangle indicates the overall sample mean effect estimated in Table 2, Column 5. Controls are as in Column 5 of Table 2. Standard errors clustered at the city level.*

Figure 6: Treatment Effects and Initial AOD levels



*Note: Graph shows the 58 city-specific subway effects as small circles based on Equation (4) where the coefficient  $\alpha$  is allowed to vary by city. Dark dots are significant at the 5% level. The x-axis shows initial AOD level in 2000, the y-axis shows coefficient estimates. The solid line shows a linear fit of the dots and the vertical dashed line shows the partition of the sample into below and above median initial AOD level. Standard errors clustered at the city level.*

## Tables

Table 1: Mean AOD and Mean Pixel Counts for New Subway Cities

	World	Africa	Asia	Europe	N. America	S. America
New Subway Cities	58	1	39	9	3	6
<u>2017</u>						
AOD, Aqua, 10km disk	0.40	0.19	0.50	0.17	0.22	0.26
AOD, Terra, 10km disk	0.41	0.20	0.50	0.18	0.20	0.24
# pixels, Aqua, 10km disk	86.95	206.20	81.74	136.68	38.93	50.27
# pixels, Terra, 10km disk	100.77	216.66	97.68	153.22	47.01	44.82
AOD, Aqua, 25km disk	0.39	0.18	0.49	0.15	0.17	0.21
AOD, Terra, 25km disk	0.39	0.19	0.49	0.17	0.18	0.23
# pixels, Aqua, 25km disk	782.62	1905.79	691.68	1236.55	614.42	597.95
# pixels, Terra, 25km disk	849.67	1899.85	791.80	1310.07	548.60	512.17
AOD, Terra, 10km footprint	0.41	0.20	0.51	0.18	0.23	0.22
# pixels, Terra, 10km footprint	87.02	150.77	87.24	129.28	38.02	29.49
AOD, Terra, 25km footprint	0.42	0.20	0.51	0.17	0.20	0.26
# pixels, Terra, 25km footprint	371.72	846.38	368.19	585.83	154.85	102.84
<u>2000</u>						
AOD, Terra, 10km disk	0.42	0.24	0.51	0.25	0.23	0.26
AOD, Terra, 25km disk	0.41	0.23	0.51	0.23	0.19	0.24
AOD, Terra, 10km footprint	0.43	0.23	0.52	0.25	0.27	0.27
AOD, Terra, 25km footprint	0.43	0.21	0.52	0.23	0.21	0.27

Table 2: Average Effect of Subway Openings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
post	-0.0128 (0.0115)	-0.0081 (0.0123)	-0.0088 (0.0089)	-0.0094 (0.0090)	-0.0094 (0.0090)	-0.0092 (0.0079)	0.0021 (0.0151)
post $\times$ t							-0.0012 (0.0014)
satellite	N	Y	Y	Y	Y	Y	Y
cont. $\times$ year	N	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	N	N	Y	Y	Y	Y	Y
climate	N	N	N	Y	N	N	N
climate $\times$ cont.	N	N	N	N	Y	Y	Y
city trend	N	N	N	N	N	N	Y
Mean AOD	0.46	0.46	0.46	0.46	0.46	0.42	0.46
$R^2$	0.61	0.62	0.80	0.80	0.80	0.76	0.80
# events	58	58	58	58	58	58	58
# cities	58	58	58	58	58	519	58
N	21806	21806	21806	21806	21806	194458	21806

Note: Dependent variable is mean AOD in a 10km disk with centroid in the city center. Column 6 adds all non-subway cities. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table 3: Heterogeneous Treatment Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
post	-0.0178* (0.0102)	-0.0087 (0.0088)	-0.0020 (0.0075)	-0.0102 (0.0087)	0.0005 (0.0088)	-0.0160 (0.0136)	-0.0081 (0.0143)	0.0032 (0.0145)	0.0126 (0.0083)
post $\times$ x	0.0169 (0.0164)	-0.0014 (0.0173)	-0.0140 (0.0162)	0.0016 (0.0167)	-0.0232 (0.0181)	0.0128 (0.0157)	-0.0026 (0.0160)	-0.0243 (0.0168)	-0.0403*** (0.0148)
satellite	Y	Y	Y	Y	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean AOD	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46	0.46
R <sup>2</sup>	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
# events	58	58	58	58	58	58	58	58	58
# cities	58	58	58	58	58	58	58	58	58
N	21806	21806	21806	21806	21806	21806	21806	21806	21806

Note: Dependent variable is mean AOD in a 10km disk with centroid in the city center. Each column uses the same specification as in Column 5 of Table 2, with the addition of an interaction term. Interaction terms x are as follows: (1) Above median city population in 2000 (2) Above median density in 2000 (3) Above median population growth, 1990-2000 (4) Below median country GDP per capita in 2000 (5) China (6) Above median elevation range in a 25km disk centered around the city center (7) Above median wind speed in 2000 (8) Above median number of stations per capita at opening (9) Above median AOD in 2000. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table 4: Effect of Subway Openings in High AOD Cities

	(1)	(2)	(3)	(4)	(5)	(6)
post	-0.0290 (0.0207)	-0.0246 (0.0212)	-0.0246 (0.0147)	-0.0272* (0.0146)	-0.0270* (0.0147)	-0.0284** (0.0130)
satellite	N	Y	Y	Y	Y	Y
cont. $\times$ year	N	Y	Y	Y	Y	Y
city $\times$ cal. mo.	N	N	Y	Y	Y	Y
climate $\times$ cont.	N	N	N	N	Y	Y
Mean AOD	0.66	0.66	0.66	0.66	0.66	0.43
bootstrap p-value	0.174	0.252	0.111	0.078	0.083	0.049
$R^2$	0.34	0.35	0.65	0.66	0.66	0.75
# events	29	29	29	29	29	29
# cities	29	29	29	29	29	490
N	10896	10896	10896	10896	10896	183548

*Note: Dependent variable is mean AOD in a 10km disk with centroid in the city center. Column 6 includes all non-subway cities. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.*

Table 5: Long Run Effects of Subway Openings in High AOD Cities

	(1)	(2)	(3)	(4)	(5)	(6)
Panel a.						
1-12 months post	-0.0177 (0.0163)	-0.0208 (0.0153)	-0.0227 (0.0160)	-0.0236 (0.0160)	-0.0304* (0.0158)	-0.0306* (0.0168)
13-24 months post	-0.0184 (0.0153)	-0.0224* (0.0130)	-0.0253 (0.0169)	-0.0279** (0.0141)	-0.0209 (0.0167)	-0.0242 (0.0150)
25-36 months post			-0.0429** (0.0161)	-0.0441*** (0.0125)	-0.0405** (0.0152)	-0.0437*** (0.0132)
37-48 months post					-0.0333 (0.0195)	-0.0416** (0.0167)
satellite	Y	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y	Y
Mean AOD	0.64	0.42	0.64	0.42	0.60	0.42
$R^2$	0.70	0.75	0.70	0.75	0.73	0.75
# events	27	27	24	24	21	21
# cities	27	488	24	485	21	482
N	10169	182821	8993	181645	7840	180492
Panel b.						
average post	-0.0180 (0.0139)	-0.0216* (0.0122)	-0.0295** (0.0132)	-0.0319*** (0.0106)	-0.0306** (0.0125)	-0.0350*** (0.0116)
bootstrap p-value	0.213	0.095	0.030	0.008	0.025	0.008

Note: Dependent variable is mean AOD in a 10km disk with centroid in the city center. Even-numbered columns include all non-subway cities. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table 6: Spatial Decay of Effects of Subway Openings in High AOD Cities

	(1)	(2)	(3)	(4)	(5)	(6)
post	-0.0270*	-0.0284**	-0.0219	-0.0272**	-0.0137	-0.0172
	(0.0147)	(0.0130)	(0.0141)	(0.0110)	(0.0151)	(0.0111)
satellite	Y	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y	Y
Mean AOD	0.66	0.43	0.64	0.40	0.59	0.38
bootstrap p-value	0.083	0.049	0.136	0.026	0.364	0.143
$R^2$	0.66	0.75	0.73	0.81	0.73	0.81
# events	29	29	29	29	29	29
# cities	29	490	29	490	29	490
N	10896	183548	10896	183446	10893	183427

*Note: Dependent variable is mean AOD. Columns 1 and 2 replicate columns 5 and 6 from Table 4 for reference. Dependent variable in columns 3 and 4 is AOD in a donut 10-25km from the city center. Columns 5 and 6 use AOD in a donut 25-50km from the city center. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.*

Table 7: Expansions

	(1)	(2)	(3)	(4)
post	-0.0044 (0.0039)	-0.0154* (0.0091)	-0.0002 (0.0031)	-0.0046 (0.0102)
post $\times$ x			-0.0076 (0.0072)	-0.0134 (0.0141)
satellite	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y
Mean AOD	0.44	0.63	0.44	0.63
$R^2$	0.78	0.73	0.78	0.73
# events	143	42	143	42
# cities	70	20	70	20
N	52268	15595	52268	15595

Note: Dependent variable is mean AOD in a 10km disk with centroid in the city center. Column (1) includes all expansions anywhere in the world (includes old subway cities outside of our main sample). Column (2) includes expansions restricted to the set of new subway cities used in Table 2. Column (3) is the same as (1) but includes the interaction of treatment with indicator for AOD in 2000 above 0.36 (median AOD in 2000 in the sample of new subway cities used in Table 2). Column (4) is the same as (2) but includes the interaction of treatment with indicator for AOD in 2000 above .36 (median AOD in 2000 in the sample of new subway cities used in Table 2). All specifications control for event fixed effects, event-specific pre-window indicators, event-specific post-window indicators, and event-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the event level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table 8: Ridership

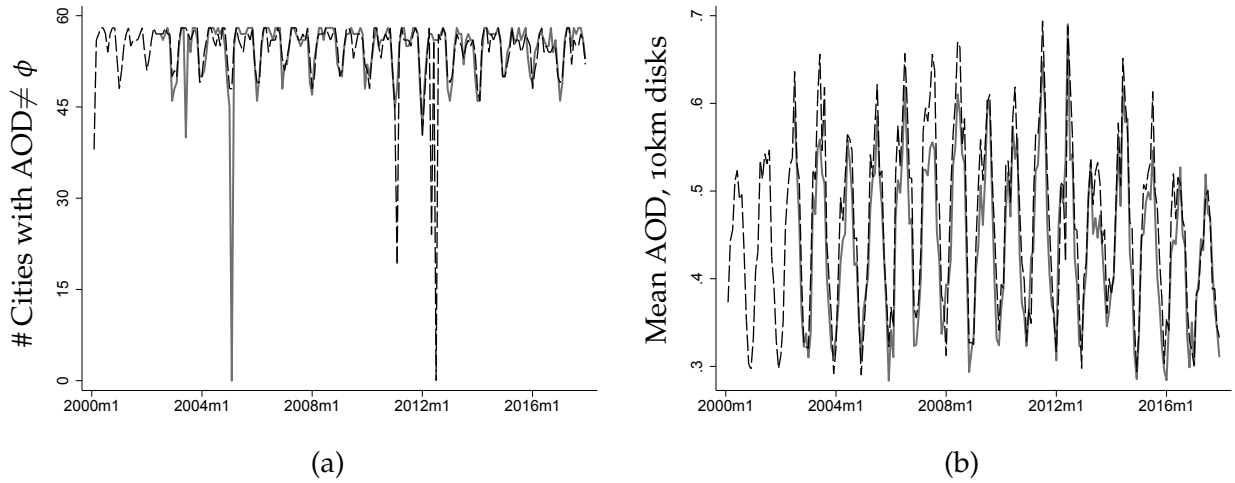
	(1)	(2)	(3)	(4)
post	36.8857*** (5.3940)	-0.0182* (0.0094)	0.0043 (0.0112)	0.0033 (0.0089)
post $\times$ x			-0.0439** (0.0163)	-0.0411** (0.0164)
satellite		Y	Y	Y
cont. $\times$ year		Y	Y	Y
city $\times$ cal. mo.		Y	Y	Y
climate $\times$ cont.		Y	Y	Y
Mean AOD		0.46	0.46	0.46
R <sup>2</sup>	0.85	0.81	0.81	0.81
# events	42	42	42	42
# cities	42	42	42	42
N	8415	15863	15863	15863

Note: (1) Dependent variable is ridership per 1000 people. (2) Dependent variable is mean AOD in a 10km disk with centroid in the city center. Controls are as in Table 2 Column 5, but restricted to sample with ridership data. (3) Same as (2) but includes interaction of treatment with indicator for above median ridership per 1000 people 12m after opening. (4) Same as (2) but includes interaction of treatment with indicator for above median ridership 12m after opening. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

## Appendix

### A AOD data

Figure A.1: MODIS Terra and Aqua AOD data



Note: Panel (a) gives count of new subway cities in our primary estimation sample with non-missing 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 2017 and Aqua for July 2002-December 2017 to construct monthly average AOD levels in cities. We download all the files from the NASA File Transfer Protocol.<sup>20</sup>

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\_3K* and *MYD04\_3K* to get AOD measures at a spatial resolution (pixel size) 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.<sup>21</sup> We

<sup>20</sup><http://ladsweb.nascom.nasa.gov/allData/6/>

<sup>21</sup>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>

consolidate 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. 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 the figure gives the count of cities in our primary estimation sample 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 in 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.

## **B Ridership data**

We gathered subway ridership data (unlinked trips) for 42 of the subway systems in our main estimating sample, mostly from annual reports or statistical agencies. In 16 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 A.1.

Table A.1: Ridership Data Sources

City	Source
Almaty (Kazakhstan)	International Metro Association reports
Bangalore (India)	Bangalore Metro operational performance
Brescia (Italy)	Brescia Mobilità reports
Changsha (China)	China Association of Metros annual reports
Chennai (India)	Chennai Metro Rail Limited annual reports
Copenhagen (Denmark)	Statistics Denmark
Daejeon (South Korea)	Daejeon Metropolitan Rapid Transit Corporation
Delhi (India)	Delhi Metro Rail Corporation annual reports
Dongguan (China)	China Association of Metros annual reports
Dubai (UAE)	Dubai Road and Transport Auth.: Annual statistical reports
Fuzhou (China)	China Association of Metros annual reports
Gwangju (South Korea)	Gwangju Subway reports
Hangzhou (China)	Hangzhou Statistical Yearbook
Harbin (China)	China Association of Metros annual reports
Jaipur (India)	Jaipur Metro annual reports
Kazan (Russia)	International Metro Association reports
Kaohsiung (Taiwan)	Kaohsiung Rapid Transit Corp. transport volume statistics
Lausanne (Switzerland)	Transports Lausanne Annual Reports
Lima (Peru)	Ministerio de Transportes y Comunicaciones Perú
Mashhad (Iran)	Mashhad Urban Railway Corp. planning and development
Mumbai (India)	Mumbai Metro One Pvt. Ltd. right to information request
Nanning (China)	China Association of Metros annual reports
Naha (Japan)	Japan Ministry of Land, Infrastructure, Transport and Tourism
Nanchang (China)	China Association of Metros annual reports
Ningbo (China)	China Association of Metros annual reports
Palma (Spain)	Instituto Nacional de Estadística España
Panama City (Panama)	Instituto Nacional de Estadísticas y Censo Panamá
Porto (Portugal)	Statistics Portugal, Light rail (metro) survey
Qingdao (China)	China Association of Metros annual reports
Salvador da Bahia (Brazil)	Companhia de Transportes do Estado da Bahia
San Juan Puerto Rico (USA)	Instituto de Estadísticas de Puerto Rico
Santo Domingo (DR)	Oficina para el Reordenamiento del Transporte
Seattle (USA)	Sound Transit performance reports (Only Central Link Line)
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
Turin (Italy)	Gruppo Torinese Transporti reports
Valparaíso (Chile)	Memoria Anual Metro Valparaíso
Wuxi (China)	China Association of Metros annual reports
Xi'an, Shaanxi (China)	Xian Bureau of Statistics
Zhengzhou (China)	China Association of Metros annual reports

We were not able to obtain ridership data from the time of opening for the following 16 cities in the sample: Algiers (Algeria), Brasilia (Brazil), Bursa (Turkey), Chengdu (China), Chongqing (China), Dalian (China), Isfahan (Iran), Izmir (Turkey), Kunming (China), Maracaibo (Venezuela), Nanjing (China), Rennes (France), Shiraz (Iran), Tabriz (Iran), Valencia (Venezuela), and Wuhan (China).

### C AOD vs ground based measurements

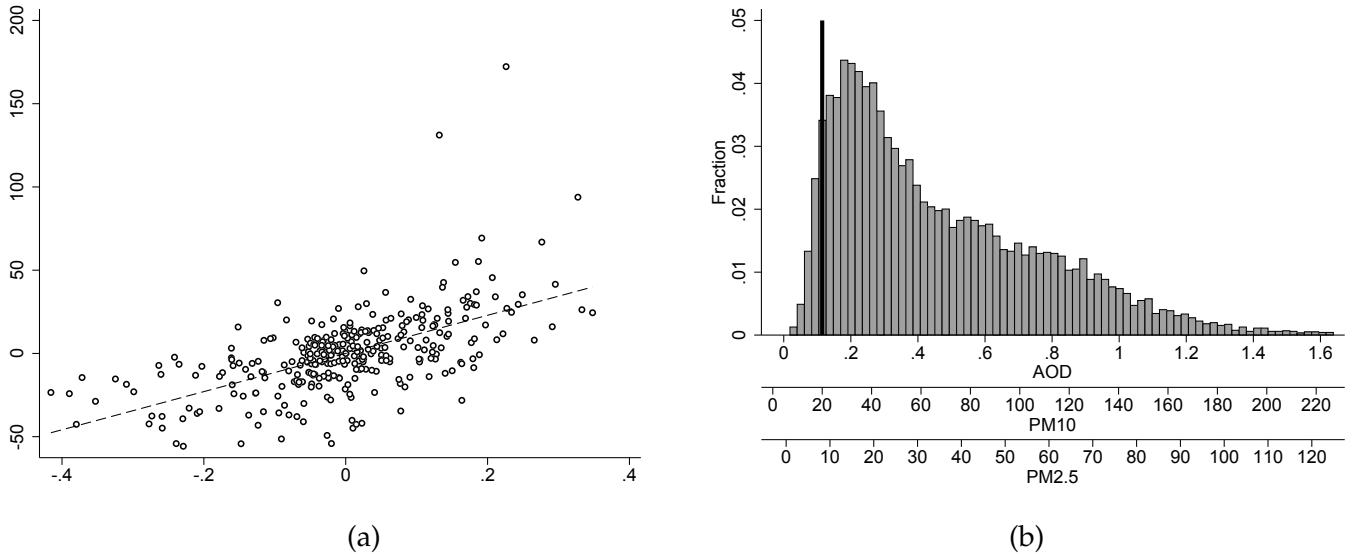
A series of papers have compared measures of AOD to measures of particulate concentration from surface instruments (e.g. Gupta, Christopher, Wang, Gehrig, Lee, and Kumar, 2006, Kumar, Chu,

Table A.2: Relationship between AOD and Ground-based Particulate Measures

	PM10						PM2.5	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AOD	135.811*** (9.687)	114.604*** (11.804)	118.094*** (11.501)	113.008*** (11.882)	101.725*** (9.982)	103.271*** (10.949)	76.573*** (6.675)	60.572*** (8.288)
cons	4.527 (2.821)	134.522*** (39.287)	110.657*** (39.842)	136.867*** (39.505)	136.647*** (41.110)	139.975*** (42.324)	-0.547 (1.546)	19.931 (27.470)
Mean dep. var.	57.25	57.25	57.25	57.25	57.35	57.35	23.72	23.72
Mean ind. var.	0.39	0.39	0.36	0.39	0.37	0.39	0.32	0.32
R <sup>2</sup>	0.50	0.82	0.81	0.82	0.83	0.82	0.61	0.85
N	340	340	340	340	339	339	217	217

Note: (1) Terra 10k disk, no controls. (2) Terra 10k disk, controls. (3) Aqua 10k disk, controls. (4) Terra 10k footprint, controls. (5) Terra 25k disk, controls. (6) Terra 25k footprint, controls. (7) Terra 10k disk, no controls. (8) Terra 10k disk, controls. Controls: continent-year indicators, average pixel count, and linear and quadratic terms in average temperature, precipitation, cloud cover, vapor pressure and frost days. Robust standard errors in parentheses. Stars denote significance levels: \* 0.10 \*\* 0.05 \*\*\* 0.01.

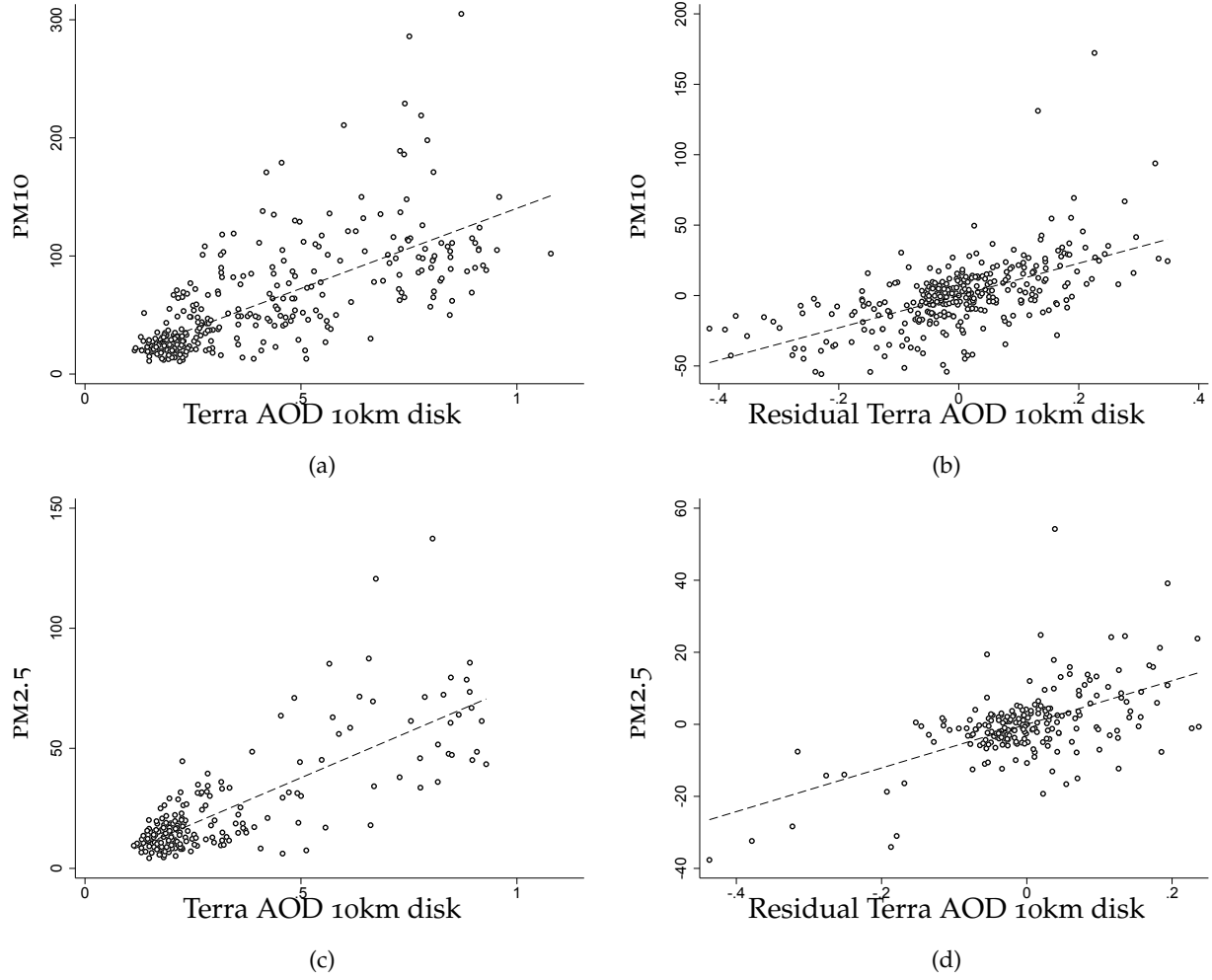
Figure A.2: AOD versus PM



Note: (a) Plot showing residualized PM<sub>10</sub> and AOD, together with linear trend. (b) Histogram of city-months by AOD, PM<sub>10</sub> and PM<sub>2.5</sub>. PM<sub>10</sub> and PM<sub>2.5</sub> axes rescaled from AOD using columns 1 and 4 of Table A.2. Black vertical line indicates WHO threshold level for annual average PM<sub>10</sub> exposure (WHO, 2006).

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. 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 loca-

Figure A.3: 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.*

tion over a longer period. This naturally causes satellite and ground based measures to diverge. Second, ground based instruments report the concentration of dry particulates, while the satellite based measure has trouble distinguishing water vapor from other particles. This motivates the use of flexible climate controls in our analysis.

As a direct check on our AOD data, we use World Health Organization data (WHO, 2016a) describing average annual PM10 and PM2.5 concentrations ( $\mu\text{g}/\text{m}^3$ ) in cities where ground-based pollution readings were available. We successfully match 150 such cities with ground-based PM10 readings to our subway cities data. Of the 150, 79 report PM10 readings during three years, 32 during two years, and 39 in only one year. The readings span the 2007-2014 period, and not all city-years record both PM10 and PM2.5. Averaging monthly AOD values to calculate yearly averages, we obtain 340 comparable city-years for PM10 and 217 comparable city-years for PM2.5.

Note the limited amount of data available from ground-based instruments. Satellite data solve the issue of data scarcity. In particular, not that the ground based measurements are annual, as opposed to the monthly data we use elsewhere.

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

$$PM_{yit} = \alpha_0 + \alpha_1 AOD_{it} + \text{controls}_{it} + \epsilon_{it},$$

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 A.2 reports results. The upper first column presents the results of a regression of the WHO measure of  $PM_{10}$  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 0.50. The AOD coefficient of 135.81 in Column 1 means that a one unit increase in AOD maps to a  $135.81 \mu g/m^3$  increase in  $PM_{10}$ . From Table 1, we see that Terra 10k readings for North America decreased by 0.03 in subway cities between 2000 and 2017. Multiplying by 135.81 gives a  $4.1 \mu g^3$  decrease in  $PM_{10}$ . By contrast, according to US EPA historical data, during this same period US average  $PM_{10}$  declined from 64.7 to  $57.7 \mu g/m^3$ , or about a 7 unit decrease.<sup>22</sup> Since Table 1 reports AOD for just the three cities in North America with new subways, while the EPA reports area weighted measures for the US, this seems as close as could be expected.

In Column 2, we conduct the same regression but include linear and quadratic terms in our climate variables, average pixel count, and continent-year indicators. The coefficient on AOD drops from 135.81 to 114.60, and the  $R^2$  increases to 0.82. In Column 3, we conduct exactly the same regression, but rely on AOD measurements from the Aqua satellite. As expected this leaves our estimates qualitatively unchanged.

Columns (4)-(6) repeat (2) but use different geographies to construct the AOD measure. In Column (4) we measure AOD in the intersection of the lights based city footprint and a 10km disk centered on the city. In Column (5) we measure AOD in a 25km disk centered on the city. In Column (6) we measure AOD in the intersection of the lights based footprint and the 25km disk. Coefficients vary slightly over the different specifications, but  $R^2$ s do not. In theory, this sequence of regressions could have revealed that ground based instruments are more closely related to a particular AOD measure. In fact, this seems not to be the case. Thus, the comparison of remotely sensed and ground based measures does not suggest that footprints are to be preferred to disks for the analysis based on  $R^2$ s. Given this, and given that most subway systems concentrate their service in the center part of the city (Gonzalez-Navarro and Turner, 2018), we rely on AOD calculated of centrally located 10km disks for our main analysis.

For completeness, columns (7) and (8) replicate columns (1) and (2), but use ground based measures of  $PM_{2.5}$  as the dependent variable. Since  $PM_{2.5}$  comprises a smaller fraction of all

<sup>22</sup><https://www.epa.gov/air-trends/particulate-matter-pm10-trends>, accessed July 2, 2020.

airborne particulates than does PM<sub>10</sub>, the smaller coefficients in these regressions is expected. In fact, the AOD coefficient for PM<sub>2.5</sub> from Column (8) is about 53% of the one for PM<sub>10</sub> in Column (2). This is consistent with the PM<sub>10</sub> to PM<sub>2.5</sub> conversion factors used by the World Health Organisation (WHO, 2016a).<sup>23</sup>

All specifications reported in Table A.2 assume a linear relationship between AOD and PM. Figure A.2(a) plots residuals of regressions of PM<sub>10</sub> and AOD on all controls used in Column 2 of Table A.2, along with a linear regression line. This graph illustrates both how closely the two variables track each other and how close to linear is the relationship between them.

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 agree seems remarkable.

In addition to validating the use of remotely sensed AOD, Table A.2 provides a basis for translating our estimates of the relationship between subways and AOD into a relationship between subways and PM<sub>2.5</sub>, or PM<sub>10</sub>. To illustrate this process, and to help to describe our data, Figure A.2(b) provides a histogram of the 21,806 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 PM<sub>10</sub> and PM<sub>2.5</sub> based on columns (1) and (7) of Table A.2. For reference, the black line in the figure gives the World Health Organization recommended maximum annual average PM<sub>10</sub> exposure level (20 µg/m<sup>3</sup>).

## D 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 PM<sub>2.5</sub> exposure, relative to an environment where PM<sub>2.5</sub> 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 PM<sub>2.5</sub> concentration  $c$  is given by  $RR_d(c, \bar{c}) = P(D_d | c) / P(D_d | \bar{c})$ , 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  $\bar{c}$

<sup>23</sup>We note that the results in Table A.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

$$\ln(\text{PM}_{2.5}) \approx 0.8 + 0.7 \ln(\text{AOD}).$$

Comparing with Table A.2, we see that these coefficient estimates are quite different. The difference reflects primarily our use of the level of PM<sub>2.5</sub>, 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 A.2 to that in Brauer *et al.* (2015) for three reasons. First, AOD is already a logarithm (see footnote 4), 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.

assumed to lie uniformly between  $5.8$  and  $8.8 \mu\text{g}/\text{m}^3$ . 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  $\text{PM}_{2.5}$  concentrations using the regression specification in Column 8 of Table A.2. Specifically, we obtain predicted  $\text{PM}_{2.5}$  values from the annual city average of AOD (and all other covariates) during the 12 months preceding the subway opening. The post-subway  $\text{PM}_{2.5}$  concentrations are obtained by subtracting  $0.028 \times 60.57 = 1.696 \mu\text{g}/\text{m}^3$  to the pre-subway concentration, where  $0.028$  is the subway AOD effect from Table 4, and  $60.57$  is the AOD coefficient in Column 8 of Table A.2.

Let  $c_0$  and  $c_1$  respectively denote the pre and post-subway  $\text{PM}_{2.5}$  concentrations in a given city. For the purpose of our burden of disease calculations, the relevant risk ratio is  $RR_d(c_1, c_0) = P(D_d | c_1) / P(D_d | c_0)$ . 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  $PAF_d = 1 - RR_d(c_1, c_0) = 1 - P(D_d | c_1) / P(D_d | c_0)$ . Interpreting  $P(D_d | c)$  as the fraction of the total population that died of disease  $d$  when exposed to  $\text{PM}_{2.5}$  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 high AOD 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  $\sum_d \sum_a PAF_{ida} \cdot M_{ida}$ .

## E Supplemental Tables

Table A.3: Average Effect of Subway Openings: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
post	-0.0094 (0.0090)	-0.0035 (0.0085)	-0.0093 (0.0090)	-0.0080 (0.0088)	-0.0050 (0.0088)	-0.0082 (0.0086)
satellite	Y	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y	Y
Mean AOD	0.46	0.41	0.45	0.46	0.45	0.47
$R^2$	0.80	0.87	0.84	0.79	0.80	0.80
# events	58	58	58	58	58	58
# cities	58	58	58	58	58	58
N	21806	21806	19635	21702	22605	22497

*Note: Dependent variable is mean AOD in a 10km disk with centroid in the city center unless noted otherwise below. (1) Replicates Column 5 from Table 2 (for reference). (2) Same as (1) but weight observations by pixel count. (3) Same as (1) but drop observations with low pixel count. (4) Same as (1) but calculating AOD in city footprint within 10km from the city center. (5) Same as (1) but calculating AOD in a 25km disk centered around the city center. (6) Same as (1) but calculating AOD in city footprint within 25km from the city center. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.*

Table A.4: Average Effect of Subway Openings: Different Window of Analysis

	(1)	(2)	(3)	(4)	(5)
post	-0.0094 (0.0090)	-0.0027 (0.0108)	-0.0040 (0.0096)	-0.0008 (0.0084)	-0.0057 (0.0088)
satellite	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y
Mean AOD	0.46	0.46	0.46	0.44	0.45
$R^2$	0.80	0.80	0.81	0.81	0.81
# events	58	64	60	55	44
# cities	58	64	60	55	44
N	21806	24028	22580	20684	16422

Note: Dependent variable is mean AOD in a 10km disk centered around the city center. (1) Column 5, Table 2 (for reference). (2) Same as (1) but treatment and control window are 6 months. (3) Same as (1) but treatment and control window are 12 months. (4) Same as (1) but treatment and control window are 24 months. (5) Same as (1) but treatment and control window are 36 months. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-0 indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table A.5: Effect of Subway Openings in High AOD Cities: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
post	-0.0295*	-0.0320**	-0.0305**	-0.0321**	-0.0258*	-0.0254**	-0.0134	-0.0238**	-0.0219	-0.0286***
	(0.0157)	(0.0134)	(0.0145)	(0.0129)	(0.0142)	(0.0124)	(0.0152)	(0.0110)	(0.0147)	(0.0108)
satellite	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean AOD	0.63	0.43	0.66	0.42	0.67	0.43	0.66	0.41	0.68	0.43
bootstrap p-value	0.102	0.028	0.046	0.028	0.091	0.049	0.389	0.045	0.146	0.014
R <sup>2</sup>	0.77	0.84	0.70	0.79	0.65	0.74	0.63	0.79	0.66	0.77
# events	29	29	29	29	29	29	29	29	29	29
# cities	29	490	29	489	29	486	29	491	29	487
N	10896	183548	9820	165235	10894	179065	11411	190020	11361	184165

Note: Dependent variable is mean AOD in a 10km disk centered around the city center unless noted otherwise below. Columns (1-2) replicate columns 5 and 6 in Table 4, but weight observations by pixel count. (3-4) Same as Columns 5 and 6 in Table 4, but drop observations with low pixel count. (5-6) Same as Columns 5 and 6 in Table 4, but calculating AOD in city footprint within 10km from the city center. (7-8) Same as Columns 5 and 6 in Table 4, but calculating AOD in a 25km disk centered around the city center. (9-10) Same as Columns 5 and 6 in Table 4, but calculating AOD in city footprint within 25km from the city center. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-o indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table A.6: Average Effect of Subway Openings in high AOD cities: Different Window Widths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
post	-0.0270* (0.0147)	-0.0284** (0.0130)	-0.0032 (0.0212)	0.0012 (0.0215)	-0.0144 (0.0164)	-0.0121 (0.0158)	-0.0116 (0.0131)	-0.0160 (0.0109)	-0.0185 (0.0136)	-0.0215** (0.0099)
satellite	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
cont. $\times$ year	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
city $\times$ cal. mo.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
climate $\times$ cont.	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean AOD	0.66	0.43	0.66	0.43	0.66	0.43	0.64	0.42	0.66	0.42
bootstrap p-value	0.083	0.049	0.863	0.952	0.374	0.436	0.381	0.150	0.170	0.046
R <sup>2</sup>	0.66	0.75	0.66	0.75	0.66	0.75	0.70	0.75	0.68	0.75
# events	29	29	29	29	29	29	27	27	22	22
# cities	29	490	29	490	29	490	27	488	22	483
N	10896	183548	10896	183548	10896	183548	10169	182821	8201	180853

Note: Dependent variable is mean AOD in a 10km disk centered around the city center. (1-2) Same as Columns 5 and 6 in Table 4 (for reference). (3-4) Same as (1-2) but treatment and control window are 6 months. (5-6) Same as (1-2) but treatment and control window are 12 months. (7-8) Same as (1-2) but treatment and control window are 24 months. (9-10) Same as (1-2) but treatment and control window are 36 months. All specifications control for city fixed effects, city-specific pre-window indicators, city-specific post-window indicators, and city-specific period-o indicators. Climate controls are pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days. Standard errors clustered at the city level in parentheses. Stars denote significance levels: \* 0.10, \*\* 0.05, \*\*\* 0.01.

Table A.7: City Level Descriptive Statistics

City	Plan approved	Construction begins	Opening date	Stations added			Daily ridership	Mean AOD	SD AOD	City population	Country GDP PC
				opening	1st exp.	2nd exp.					
Rennes (France)	1989	Jan. 1997	Mar. 2002	15	n.a.	n.a.	n.a.	0.20	0.10	283	33,534
Bursa (Turkey)	1997	Jan. 1998	Aug. 2002	17	n.a.	n.a.	n.a.	0.26	0.09	1,290	10,544
Copenhagen (Denmark)	1995	Nov. 1996	Nov. 2002	11	n.a.	n.a.	87,811	0.16	0.08	1,089	34,706
Porto (Portugal)	1996	Mar. 1999	Dec. 2002	8	6	7	35,200	0.18	0.08	1,264	23,297
Delhi (India)	1995	Jan. 1998	Dec. 2002	6	3	22	184,000	0.74	0.29	16,956	2,144
Dalian (China)	1999	Sep. 2000	May 2003	12	6	17	n.a.	0.54	0.23	3,338	4,993
Naha (Japan)	1996	Nov. 1996	Aug. 2003	15	n.a.	n.a.	31,237	0.29	0.15	306	34,120
Tianjin (China)	1984	Jan. 2001	Mar. 2004	7	25	17	n.a.	0.64	0.24	7,901	5,239
Gwangju (South Korea)	1994	Aug. 1996	Apr. 2004	13	n.a.	n.a.	36,780	0.41	0.22	1,388	25,527
Wuhan (China)	1999	Dec. 2000	Sep. 2004	25	20	13	n.a.	0.87	0.28	7,036	5,461
Shenzhen (China)	1992	Dec. 1998	Dec. 2004	18	69	13	249,722	0.76	0.28	8,087	5,572
San Juan Puerto Rico (USA)	1992	Jul. 1996	Apr. 2005	16	n.a.	n.a.	24,000	0.22	0.11	2,495	49,527
Chongqing (China)	1983	Dec. 2000	Jun. 2005	13	12	16	n.a.	0.91	0.25	9,293	5,900
Kazan (Russia)	1989	Aug. 1997	Aug. 2005	5	n.a.	n.a.	53,055	0.22	0.15	1,117	13,581
Nanjing (China)	1994	Dec. 2000	Aug. 2005	13	3	29	n.a.	0.77	0.29	5,076	6,010
Valparaiso (Chile)	1999	May 2002	Nov. 2005	20	n.a.	n.a.	32,755	0.10	0.03	842	13,324
Turin (Italy)	1999	Dec. 2000	Feb. 2006	11	n.a.	n.a.	38,275	0.26	0.13	1,707	33,214
Daejeon (South Korea)	1996	Jan. 1996	Mar. 2006	12	n.a.	n.a.	68,939	0.38	0.23	1,437	27,555
Valencia (Venezuela)	1994	Nov. 1997	Nov. 2006	3	n.a.	n.a.	n.a.	0.23	0.11	1,548	11,531
Maracaibo (Venezuela)	1998	Mar. 2004	Nov. 2006	2	n.a.	n.a.	n.a.	0.28	0.13	1,707	11,531
Kaohsiung (Taiwan)	1994	Jan. 2001	Mar. 2008	23	13	n.a.	122,243	0.54	0.21	1,508	38,122
Palma (Spain)	2004	Aug. 2005	Jul. 2008	9	6	n.a.	4,374	0.18	0.06	375	34,865
Lausanne (Switzerland)	2000	Feb. 2004	Oct. 2008	14	n.a.	n.a.	66,852	0.19	0.07	356	54,124
Santo Domingo (DR)	2005	Nov. 2005	Jan. 2009	16	13	n.a.	62,293	0.30	0.15	2,480	10,075

Continued on next page

Table A.7 – continued from previous page

City	Plan approved	Construction begins	Opening date	Stations added			Daily ridership	Mean AOD	SD AOD	City population	Country GDP PC
				opening	1st exp.	2nd exp.					
Adana (Turkey)	1988	Sep. 1996	Mar. 2009	8	n.a.	n.a.	n.a.	0.33	0.12	1,453	16,317
Seville (Spain)	1999	Aug. 2005	Apr. 2009	17	n.a.	n.a.	43,461	0.20	0.08	694	34,496
Seattle (USA)	1996	Nov. 2003	Jul. 2009	8	n.a.	n.a.	15,437	0.17	0.08	3,017	49,706
Dubai (UAE)	2005	Mar. 2006	Sep. 2009	10	16	n.a.	169,816	0.51	0.24	1,699	65,788
Chengdu (China)	2000	Dec. 2005	Sep. 2010	16	20	5	n.a.	0.90	0.32	7,481	9,131
Shenyang (China)	1999	Nov. 2005	Sep. 2010	22	19	n.a.	434,428	0.48	0.26	5,819	9,131
Xian, Shaanxi (China)	1994	Sep. 2006	Sep. 2011	17	26	n.a.	295,220	0.73	0.25	5,684	10,043
Bangalore (India)	2003	Apr. 2007	Oct. 2011	7	10	n.a.	18,436	0.47	0.12	8,579	4,461
Mashhad (Iran)	1994	Dec. 1999	Oct. 2011	21	6	n.a.	83,345	0.22	0.10	2,717	18,195
Algiers (Algeria)	1988	Mar. 1993	Nov. 2011	10	n.a.	n.a.	n.a.	0.33	0.17	2,461	13,342
Almaty (Kazakhstan)	1980	Sep. 1988	Dec. 2011	7	n.a.	n.a.	18,228	0.25	0.06	1,470	22,022
Lima (Peru)	1986	Oct. 1986	Apr. 2012	16	n.a.	n.a.	120,575	0.71	0.19	9,150	10,496
Suzhou, Jiangsu (China)	2002	Dec. 2007	Apr. 2012	24	20	38	196,778	0.81	0.25	4,326	10,365
Kunming (China)	2009	May 2010	Jun. 2012	2	12	18	n.a.	0.31	0.24	3,602	10,423
Hangzhou (China)	2005	Mar. 2007	Nov. 2012	31	13	10	390,595	0.79	0.22	6,112	10,568
Brescia (Italy)	2000	Jan. 2004	Mar. 2013	17	n.a.	n.a.	42,943	0.27	0.14	452	36,113
Harbin (China)	2005	Sep. 2009	Sep. 2013	17	n.a.	n.a.	182,333	0.30	0.18	5,457	11,041
Zhengzhou (China)	2008	Jun. 2009	Dec. 2013	20	15	20	244,722	0.83	0.34	4,074	11,189
Changsha (China)	2008	Sep. 2009	Apr. 2014	18	n.a.	n.a.	233,528	0.83	0.25	3,799	11,374
Panama City (Panama)	2009	Feb. 2011	Apr. 2014	12	n.a.	n.a.	223,661	0.33	0.14	1,615	18,887
Ningbo (China)	2003	Jun. 2009	May 2014	20	21	n.a.	104,889	0.81	0.22	3,187	11,420
Mumbai (India)	2004	Feb. 2008	Jun. 2014	12	n.a.	n.a.	260,000	0.59	0.35	18,992	5,095
Salvador da Bahia (Brazil)	1999	Apr. 2000	Jun. 2014	5	2	n.a.	42,782	0.20	0.07	3,517	15,231
Wuxi (China)	2006	Nov. 2009	Jul. 2014	24	18	n.a.	229,642	0.84	0.24	2,915	11,513
Shiraz (Iran)	1993	Jun. 2006	Oct. 2014	5	n.a.	n.a.	n.a.	0.27	0.09	1,513	15,816

Continued on next page

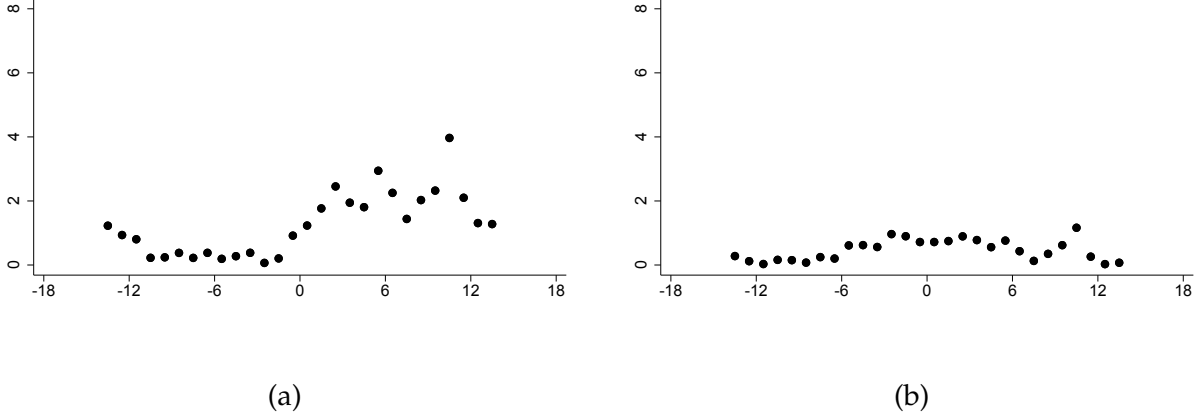
Table A.7 – continued from previous page

City	Plan approved	Construction begins	Opening date	Stations added			Daily ridership	Mean AOD	SD AOD	City population	Country GDP PC
				opening	1st exp.	2nd exp.					
Chennai (India)	2008	Jun. 2009	Jun. 2015	6	n.a.	n.a.	10,923	0.42	0.11	9,554	5,391
Jaipur (India)	2010	Feb. 2011	Jun. 2015	9	n.a.	n.a.	20,064	0.41	0.21	3,380	5,391
Tabriz (Iran)	2000	n.a.	Aug. 2015	6	n.a.	n.a.	n.a.	0.30	0.09	1,538	15,152
Isfahan (Iran)	1996	Jun. 2001	Oct. 2015	10	n.a.	n.a.	n.a.	0.40	0.10	1,901	14,994
Nanchang (China)	2006	Jul. 2009	Dec. 2015	24	16	n.a.	304,894	0.79	0.25	3,053	12,034
Qingdao(China)	2009	Nov. 2009	Dec. 2015	9	18	n.a.	182,589	0.72	0.23	5,041	12,034
Dongguan (China)	2007	Mar. 2010	May 2016	15	n.a.	n.a.	107,611	0.88	0.25	7,330	12,229
Fuzhou (China)	2009	Dec. 2009	May 2016	9	n.a.	n.a.	136,969	0.50	0.18	3,344	12,229
Nanning (China)	2008	Dec. 2011	Jun. 2016	10	17	n.a.	269,808	0.85	0.30	3,353	12,268
Average	1999	Jan. 2004	Jun. 2010	13.8	16.0	17.3	130,528	0.47	0.32	4,028	17,755

*Note: Daily ridership reported 18 months after opening. Mean and SD AOD columns report mean and standard deviation values in a 10km radius using Terra satellite monthly observations from 2000-2017. City population column reports metropolitan area population in thousands at time of subway opening. Country GDP PC reports PPP-adjusted country GDP per capita from Penn World Tables Version 9.1 reported at time of subway opening.*

## F Supplemental Figures

Figure A.4: Structural Break Tests



Notes: (a) Plot of Wald statistics for tests of a regression intercept discontinuity at time  $\tau$ . Test statistics calculated in regressions that also control for a satellite indicator, year-by-continent indicator variables, city-by-calendar month indicators, and continent specific climate variables (AOD pixel count and linear and quadratic terms in temperature, precipitation, cloud cover, vapor pressure and frost days). (b) Plot of Wald statistics for tests of a trend break at time  $\tau$  conditional on a discontinuity in the mean level of AOD at  $\tau = 1$ . Other details are the same as in Panel (a).

The regressions underlying Figure 4(a) are,

$$AOD_{it} = \beta_i + \alpha_{1j}D_{it}(j,k) + \gamma'X_{it} + \epsilon_{it} \quad (5)$$

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 (5) for each month in  $j \in \{-14, \dots, 14\}$  with errors clustered at the city level. We then calculate a Wald test of  $\alpha_{1j} = 0$  for each  $j$ . Including pre- and post-period indicators in these regressions allows us to 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 estimate the effect of subways on AOD. Here and throughout, we allow for city specific pre- and post-treatment windows. This extra flexibility helps to insure that we use the same variation in the data to estimate the subway effect, even as we change samples. We discuss this further below. Panel (a) of Figure 4 plots these test statistics.

The figure shows a clear pattern. Wald statistics increase from a low level to a peak between month 6 and 12 of subway opening. 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, much greater 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. In conclusion, while we cannot reject a null of no level break in the overall sample, it is important to note that the plot suggests that whatever pollution reducing effects the subway system opening leads to occur in the months *after* the subway opens.

In Figure 4(b) we 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 conditional on a level break at  $\tau = 1$ . Formally, this means estimating the following set of regressions,

$$AOD_{it} = \beta_i + \alpha_1 \tau_{it} + \alpha_2 j \tau_{it} D_{it}(j, k) + \alpha_3 D_{it}(1, k) + \gamma' X_{it} + \epsilon_{it} \quad (6)$$

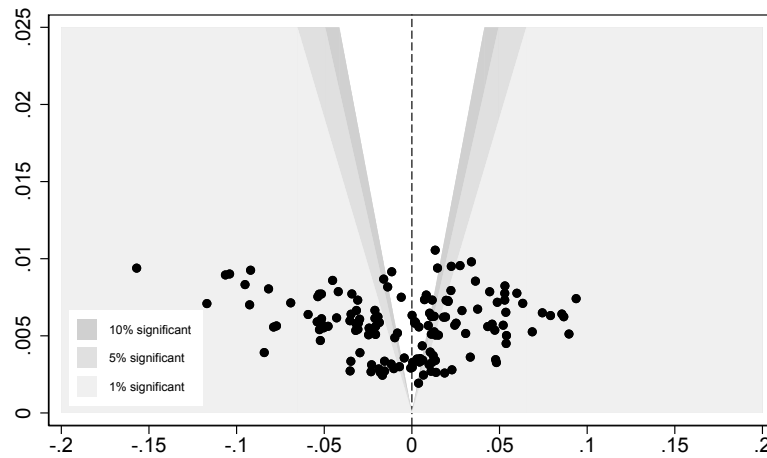
As before, we estimate the regression (6) for each month in  $j \in \{-14, \dots, 14\}$  with errors clustered at the city level and calculate the Wald test for  $\alpha_2 = 0$  for each regression.<sup>24</sup> Panel (b) of Figure 4 plots these Wald statistic values as  $j$  varies.<sup>25</sup> Thus, conditional on a step at  $\tau = 1$ , subway openings do not seem to cause a change in the trend of AOD in a city.

---

<sup>24</sup>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.

<sup>25</sup>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.

Figure A.5: Heterogeneous treatment effects for all expansions.



*Note: Illustration of all event-specific subway expansion effects. Controls are as in Column 1 of Table 7.  $x$  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. Dark, medium and light gray regions are different from zero at 10%, 5% and 1% in two-sided tests. Standard errors clustered at the event level.*