Trans-Boundary Air Pollution Spillovers: Physical Transport and Economic Costs by Distance

Fu, Shihe and Viard, Brian and Zhang, Peng

Xiamen University, Cheung Kong Graduate School of Business, The Chinese University of Hong Kong, Shenzhen

14 August 2019

Trans-Boundary Air Pollution Spillovers:
Physical Transport and Economic Costs by Distance*

Shihe Fu
Xiamen University
fushihe@xmu.edu.cn

V. Brian Viard
Cheung Kong Graduate School of Business
brianviard@ckgsb.edu.cn

Peng Zhang
The Chinese University of Hong Kong, Shenzhen
jumpersdu@gmail.com

This version: 8/15/2020

Abstract

The economic costs of trans-boundary pollution spillovers versus local effects is a necessary input in evaluating centralized versus decentralized environmental policies. Directly estimating these for air pollution is difficult because spillovers are high-frequency and vary with distance while economic outcomes are usually measured with low-frequency and local pollution is endogenous. We develop an approach to quantify local versus spillover effects as a flexible function of distance utilizing commonly-available pollution and weather data. To correct for the endogeneity of pollution, it uses a mixed two-stage least squares method that accommodates high-frequency (daily) pollution data and low-frequency (annual) outcome data. This avoids using annual pollution data which generally yields inefficient estimates. We apply the approach to estimate spillovers of particulate matter smaller than 10 micrograms (PM_{10}) on manufacturing labor productivity in China. A one $\mu g/m^3$ annual increase in PM_{10} locally reduces the average firm's annual output by CNY 45,809 while the same increase in a city 50 kilometers away decreases it by CNY 16,248. The spillovers decline quickly to CNY 2,847 at 600 kilometers and then slowly to zero at about 1,000 kilometers. The results suggest the need for supra-provincial environmental policies or Coasian prices quantified under the approach.

JEL Codes: D62; Q51; Q53; R11
Key words: air pollution; spillovers; environmental costs and benefits, mixed two-stage least squares; regional coordination

* We thank Michael Anderson, Richard Arnott, Tom Chang, Tatyana Deryugina, Andrew Foster, Qing Han, Ying Liu, Dick Morgenstern, Adriana Lleras-Muney, Ruixue Jia, Alberto Salvo, Nathan Schiff, Katja Seim, and seminar and conference participants at UC Riverside, 2019 AREUEA International Conference, Shanghai University of Finance and Economics, Nankai University, Jinan University, Wuhan University, Hunan University, 2019 Asian Meeting of Econometric Society, and National University Singapore for helpful comments. Quanyun Song, Jianghao Wang, Castiel Zhuang, and Shihan Shen provided excellent research assistance. Shihe Fu acknowledges financial support by the National Natural Science Foundation of China (Grant # 71773096 and Basic Scientific Center Project 71988101).

1. Introduction

Since the seminal work of Oates (1972) on fiscal federalism, there has been a debate on whether centralized or decentralized policies can achieve the most efficient outcome. Local authorities have better information about costs and benefits and can better tailor local policies than central authorities whose policies tend to be overly uniform. However, local jurisdictions generally ignore the effects of their policies on other jurisdictions unless these are internalized administratively. Clear and enforceable assignment of property rights followed by Coasian bargaining can also solve these externalities even under decentralized control (Coase, 1960) but require knowledge and quantification of the extra-territorial damages incurred as a function of distance.

Despite this, we are not aware of any studies that quantify trans-boundary spillovers relative to local effects for any kind of pollution. Previous papers show that trans-boundary pollution spillovers exist and that they affect extra-territorial economic well-being[1] but they do not quantify how spillovers compare to local effects as a function of distance. Our paper aims to fill this gap by providing an approach for estimating an air pollution spillover gradient including local effects for endogenous economic outcomes.

Air pollution is a prototypical example of the fiscal federalism debate with serious welfare implications. High levels of air pollution in developing countries have led to adverse effects on health, economic output, and physical and mental comfort. Ninety-two percent of all air pollution-related deaths are estimated to occur in low- and middle-income countries and ambient air pollution is estimated to have cost 4.4% of global GDP in 2016 (Ostro, *et al*., 2018). Air pollution levels far exceed the social optimum because spillovers, including trans-boundary, are not internalized. Developed countries also recognize the need to manage cross-boundary pollution to address these externalities. For example, the U.S. Clean Air Act Section 126 allows a downwind state to petition the Environmental Protection Agency to take action against an upwind state that impedes its ability to comply with smog standards.[2]

Regardless of the method used to correct the externality, a necessary input is the magnitude and geographic extent of the spillovers by distance. Centralized decision-making to internalize spillovers requires knowledge of how far spillovers extend at significant levels. Alternatively, assigning property rights and allowing for decentralized Coasian bargaining requires a method for the parties to estimate the

[1] These include Sigman (2002), Sigman (2005), Zheng *et al.* (2014), Bošković (2015), Kahn *et al.* (2015), Cai *et al.* (2016), Altindag *et al.* (2017), Jia and Ku (2017), Lipscomb and Mobarak (2017), Sheldon and Sankaran (2017), and Goodkind *et al*. (2019). We comment more on these below.

[2] Described at https://www.epa.gov/ground-level-ozone-pollution/ozone-national-ambient-air-quality-standards-naaqs-section-126.

origin of spillovers and their damage. To estimate air pollution spillovers requires estimating not just the quantity of pollution that drifts as a function of distance but also the economic costs that it imposes upon arrival. Finally, a quantification of local effects is required to determine whether the spillovers are important in relative terms.

If pollution, weather, and outcome data are available on a daily basis estimating the effect of spillovers on the outcome is straightforward: a reduced form estimate of imported pollution on local economic outcomes. However, many economic outcomes are measured at a lower frequency (e.g., annual) and air pollution spillovers occur according to daily wind patterns. Aggregating data to the annual level and directly relating economic outcomes to imported pollution is likely to involve significant efficiency losses as we show occurs in our application. In addition, reduced-form estimates do not quantify the local causal effects. We develop an approach to overcome this and demonstrate it by estimating effects of air pollution spillovers on annual manufacturing labor productivity in China.

Our approach relies on the fact that there are two determinants of the trans-boundary effect of pollution on an outcome: how much air pollution is physically transported across cities (the pollution spillover) and the causal effect of this pollution on the outcome upon its arrival in the destination city. We wish to estimate the pollution spillover flexibly to allow for a highly nonlinear gradient. However, the causal effect requires instruments for pollution and is therefore constrained to linear estimating equations. To accomplish this, we proceed in two steps. In the first step, we estimate the pollution spillover (which we call the spillover decay function) of nearby- on focal-city pollution flexibly as a function of distance using *daily* data conditional on wind blowing toward the focal city. In the second step, we estimate the causal effect of focal-city air pollution on the economic outcome. Multiplying the spillover decay effects from the first step by the causal effect from the second step is equivalent to a reduced-form approach[3] and allows us to estimate spillovers on the outcome flexibly over a range of distances and compare them to the local effect.

When we estimate the causal effect of pollution in the second step, we instrument for the endogeneity of focal-city air pollution using the air quality of the nearest nearby city conditional on wind blowing toward the focal city. When wind blows toward the focal city, imported pollution from the nearby city degrades focal-city air quality. Although other instruments could be used in this step, using nearby-city pollution is convenient because the required data (daily pollution and wind measures) are

[3] Although the spillover decay function is estimated at the daily level, the effects can be interpreted as the annual effects of a sustained and uniform increase in nearby-city pollution on all days of the year if wind blew toward focal cities on all days. Since the wind blows toward focal cities roughly half the time on average, annual spillovers are roughly half the daily effect as we describe in our results.

commonly available and are already used to estimate the pollution decay function in the first step. The exogeneity of this instrument requires high-frequency data for two reasons. First, to capture wind direction shifts precisely enough and, second, to preclude confounding factors affecting both nearby-city pollution and focal-city economic outcomes that might occur over longer time periods (in particular inter-regional economic shocks).[4] We provide evidence that daily data are frequent enough but higher levels of aggregation are not.

To combine the daily instrumenting data with the annual outcome data, we employ mixed two-stage least squares (M2SLS) (Dhrymes and Lleras-Muney, 2006), a methodology for implementing 2SLS with different levels of aggregation in the two stages. While the daily instrumenting data can be annualized (conditional on wind direction) to use Wald 2SLS, we show in our application that this results in very inefficient estimates relative to M2SLS. This is likely to be the case in estimating the effect of pollution on other annual outcomes because of the information loss that occurs when daily data is averaged to the annual level in the first stage.

We demonstrate this approach by estimating the effect of trans-city drifts of particulate matter less than 10 micrograms in diameter (PM_{10}) on short-run manufacturing labor productivity in China using a large firm-level data set from 2001 to 2007. A one $\mu g/m^3$ annual increase in PM_{10} in a city within 50 kilometers decreases the average firm's annual labor productivity by CNY 16,248 (0.106%).[5] This effect declines quickly to CNY 2,847 (0.019%) for nearby cities at 550-600 kilometers after which it declines slowly to zero at about 1,000 kilometers compared to a local effect of CNY 45,809 (0.300%). Thus, the spillover is roughly 35.5% of the local effect at 50 kilometers, falling to 6.2% at 550 kilometers, and zero at 1,000 kilometers and beyond. While we demonstrate the estimation approach with PM_{10} and productivity, it can be easily tailored to estimate the spillovers for other pollutants and other annual outcomes such as GDP, morbidity, and mortality.

This paper contributes to three strands of literature. First, we quantify the magnitude of spillovers as a function of distance relative to local effects, a key input in choosing centralized versus decentralized environmental policies (Oates and Schwab, 1988; Ogawa and Wildasin, 2009; Banzhaf and Chupp, 2012; Eichner and Runkel, 2012; Williams, 2012; Fell and Kaffine, 2014). Extant work on trans-boundary spillovers either shows that trans-boundary pollution spillovers exist (Sigman, 2002; Sigman, 2005; Kahn *et al.*, 2015; Cai *et al.*, 2016; Lipscomb and Mobarak, 2017) or that they affect extra-territorial economic well-being (Zheng *et al.*, 2014; Bošković, 2015;

[4] Exogeneity also requires that wind direction is random with respect to nearby-city pollution conditional on control variables. We provide evidence that this is the case.
[5] This estimate is for the average city given average weather.

Altindag *et al*., 2017; Sheldon and Sankaran, 2017; Jia and Ku 2019) but do not quantify their extensiveness or size relative to local effects.

Second, we develop an approach based on M2SLS that allows high-frequency variation in wind direction to be used as an instrument for high-frequency air pollution in estimating its causal effect on low-frequency outcomes. There are two approaches to using wind direction as an instrument. One approach is to use dominant wind direction alone without measures of non-local pollution sources (Deryugina *et al*., 2019; Freeman *et al*., 2019; Herrnstadt *et al*., 2019; Anderson, 2020). This is convenient because the instrument is valid without the need to measure non-local pollution. The downside, as Deryugina *et al*. (2019) points out, is that the monitoring stations that measure local pollution must be geographically dense enough to avoid measurement error and confounding effects from local pollution sources.[6] The second approach combines wind direction with the extra identification from non-local pollution. The advantage of this is that it is not confounded by local sources of pollution and can be used in the absence of a dense network of local monitoring stations. The downside is that non-local pollution sources must be measured and must be orthogonal to local sources. Previous papers that use this approach (Schlenker and Walker, 2016; Rangel and Vogel, 2019)[7] use discrete, exogenous events that shift non-local pollution. Our paper takes this approach but extends it to use a continuous measure of non-local pollution and allow for the instrument to be of higher frequency than the endogenous variable.

Third, our paper adds to the growing literature on estimating air pollution's effect on labor productivity (Graff Zivin and Neidell, 2012; Chang *et al*., 2016; Fu *et al*., 2018; Chang et *al*., 2019; He *et al*., 2019). These papers estimate the effect of an increase in local air pollution on local firms' productivity. In contrast to previous papers, we distinguish the effect of local and imported pollution sources on productivity and show that spillovers can contribute significantly to productivity losses.

We find that pollution exerts a substantial negative effect on productivity even at relatively far distances. Twenty-two percent of PM_{10} produced from a city within 300 kilometers is imported into a focal city when the wind blows directly toward it. From a policy perspective, to internalize this would require centralized control of

[6] As they explain, having a dense network of monitors locally averages out the effects of local pollution sources so that they do not bias estimates. Slightly modifying their example (page 14) imagine a smokestack in the middle of a city. If there is a single monitor on the east side of the city then the monitor will detect the pollution from the smokestack when the wind is blowing from the west but not when it blows from the east and the wind direction instrument is correlated with local pollution. However, if there is a dense network of monitors on all sides of the smokestack then a local pollution measure averaged across all monitors will reduce, and in the limit, eliminate this correlation.
[7] Schlenker and Walker (2016) also use wind speed which provides further variation besides wind direction to ensure exogeneity.

administrative areas that are 300 kilometers in radius or 283-thousand square kilometers. This is greater in size than many medium-sized provinces in China such as Hunan, Shaanxi, Hebei, Jilin, Hubei, and Guangdong (Ministry of Civil Affairs, 2017). Thus, the results indicate that environmental policies need to be coordinated at the supra-provincial level to internalize spillovers. The other major policy application of our method is in calculating Coasian prices as a decentralized solution to air pollution externalities. Our estimates allow a quantification of the compensation that one city must make to another to internalize inter-city pollution damage given the distance between the two cities, the annual wind-direction distribution, and annual levels of the economic outcome of interest. We provide an example in our results.

The scientific literature uses an alternative approach for the first step of our procedure, chemical-transport models or CTMs, to relate source emissions to receptor concentrations (Moussiopoulos, *et al.* (1996); Seigneur and Moran (2004); Seigneur and Dennis (2011)). CTMs that estimate this relationship over long distances such as we do are grid-based models that relate locations defined by three-dimensional grids that are normally one kilometer or larger in size.[8] The relationships are based on detailed mathematical models of atmospheric processes using detailed weather and emissions data. As an alternative for the first step of our procedure, CTMs offer advantages and disadvantages relative to our approach.

CTMs quantify the spillovers from original emissions and is unaffected by their displacement unlike our approach which relies on concentrations (hence the need for daily wind data to identify spillovers in our estimation). On the other hand, detailed emissions data are often not available while concentrations are more readily available. Relatedly, CTMs require highly disaggregated data on weather and pollution which is often not available, especially in developing economies. CTMs realistically model the processes of concentration formation and movement; however this greater complexity involves longer solutions times and many more assumptions. In a policy context, agreeing upon these assumptions can require significant effort and resources.[9] In contrast, our approach can be estimated quickly and its transparency requires agreement on fewer assumptions.

Our results have specific implications for the role of China's governance system in air pollution spillovers. China's reforms have succeeded in part because of its

[8] The other approach, known as source-specific models, identify specific emissions sources that contribute to ambient concentrations but are applicable up to only about 150 kilometers between source and receptor locations.

[9] For example, the EPA devotes significant resources in choosing which models meet their standards via conferences, technical analyses, and regulatory reports. A recent example is detailed in Federal Register (2017).

regionally decentralized system in which the central government provides incentives to local governments based primarily on local GDP to the exclusion of other criteria (Jin *et al*., 2005; Li and Zhou, 2005; Xu, 2011) such as environmental quality. Our results indicate that these incentives exacerbate the negative implications of air pollution spillovers on manufacturing productivity. This complements Jia (2017) which provides empirical evidence that these incentives result in more pollution. Including local environmental quality in local government officials' performance valuation is not enough; cross-boundary pollution spillovers must be considered too.

The remainder of the paper proceeds as follows. The next section describes the data we use to illustrate the estimation approach and Section 3 the approach. Section 4 provides the results, and Section 5 concludes.

2. Data

We estimate pollution spillovers on labor productivity for manufacturing firms in China from 2001 to 2007 in two steps. The first step (estimating the pollution decay function) requires daily pollution and weather data. The second step of the procedure (estimating the causal effect of air pollution on productivity) requires daily data for the instrument to address the endogeneity of pollution and accommodates annual data on productivity.

2.1 Pollution data

The highest-frequency pollution data available with significant geographic coverage during our sample period is the daily Air Pollution Index (API) published by the Ministry of Ecology and Environment. This is available at the city level and only for larger cities. The number of cities reporting API data increases over time in the sample. The sample includes 60 unique cities (Appendix A shows their location).

The API ranges from 0 to 500 with higher values indicating higher pollution concentrations and more harmful health effects (Andrews, 2008). During the sample period, a city's daily API reports the worst of three pollutants: particulate matter (PM_{10}), nitrogen dioxide (NO_2), and sulfur dioxide (SO_2) whose concentrations are measured at multiple monitoring stations within the city. Each is rescaled as an API measure to make them comparable and the pollutant with the maximum API is reported.[10] The identity of the maximal pollutant is reported if the API exceeds 50.

[10] Each monitoring station records the concentrations of the three pollutants multiple times a day. Each of these intra-day measurements is rescaled to an API index. A daily mean API for each pollutant across all stations in a city is then calculated and the maximum of these three means is the city-level API for that day.

The API is potentially subject to manipulation by those who collect and report the data. Using 2001 to 2010 data, Ghanem and Zhang (2014) find a discontinuity in the API distribution around 100 which suggests that self-reported data is manipulated by local officials who are evaluated on the annual number of "Blue Sky" days (those below 100). Also consistent with this, Andrews (2008) finds that a significant number of days in 2006 and 2007 with reported API values between 96 and 100 would fall in the range 101 to 105 if calculated using the underlying monitoring station data. To avoid any possible bias in the estimates we exclude days when the API is between 95 and 105 in either the focal or nearby city in the main estimates but show that it is robust to including these.

We use PM_{10} in the analysis rather than the API index because we wish to use physical pollution levels in quantifying spillovers and PM_{10} is overwhelmingly the worst of the three pollutants (about 90% of days). We drop days in which PM_{10} is not the maximal pollutant and for the remaining days infer its value from the API based on the piecewise-linear relationship between PM_{10} and the API (Appendix B). Although we do not observe the worst pollutant when the API is below 50 we assume it is PM_{10} in the baseline estimates because at these low levels air quality is assumed to be safe regardless of pollutant. The results are robust to dropping these days.

2.2 Wind and weather data

We require daily wind data for estimating the spillover decay function and to instrument pollution when estimating its effect on productivity. We use station-level wind direction data from the World Weather Records Clearinghouse collected by the U.S. National Oceanic and Atmospheric Administration (NOAA).[11] The data provide a direction from which the wind is blowing stated in degrees clockwise from true North in each three-hour period of each day in each city. We use a "unit-vector" average method defined by the NOAA to arrive at an average daily wind direction for each city.[12] For wind direction we use data for the focal not the nearby city. Regardless of the wind direction in the nearby city, pollution cannot be imported if the wind in the focal city is not blowing from the nearby city's direction.

[11] Data available at: http://www.ncdc.noaa.gov/data-access.

[12] In each three-hour period, we convert the direction for each monitoring station to a unit vector with coordinates $\langle u, v \rangle$. The u-component is the North-South wind direction and v the East-West. We average the two coordinates separately across the periods of each day and all stations to yield $\bar{u}$ and $\bar{v}$. We then translate the direction into a 0 to 360 degree scale based on the signs of $\bar{u}$ and $\bar{v}$: $180 - \theta$ if $\bar{u} < 0$ and $\bar{v} > 0$, $\theta - 180$ if $\bar{u} < 0$ and $\bar{v} < 0$, $360 - \theta$ if $\bar{u} > 0$ and $\bar{v} < 0$, and θ if $\bar{u} < 0$ and $\bar{v} > 0$ where $\theta = (180/\pi) * arctan(\bar{u}/\bar{v})$. This is method 1 described at: http://www.ndbc.noaa.gov/wndav.shtml.

To control for weather conditions that affect the transport of pollution and productivity we use daily weather (humidity, windspeed, and temperature) data downloaded from the Weather Underground.[13]

2.3 Firm productivity data

Our firm-level output and characteristics data are from annual surveys of manufacturing firms conducted by China's National Bureau of Statistics (NBS). The survey includes all state-owned enterprises (SOEs) regardless of size and all non-SOEs whose annual sales exceed CNY 5 million (USD 0.8 million).[14] The survey also contains detailed information on firm location, accounting measures, and firm characteristics. Before we match with the pollution data this captures 90.7% of China's total manufacturing output during the sample period (Brandt *et al*., 2012). We follow Brandt *et al.* (2012) in matching firms over time to form an unbalanced panel and in converting nominal into real values using industry-level price indices. To be consistent with the previous literature, we drop observations with missing or unreliable data (Cai and Liu, 2009; Brandt *et al.*, 2012; Yu, 2014) and winsorize the top and bottom 0.5% of data based on each of the values of output, value added, employment, and capital (Cai and Liu, 2009).

We measure output as value added per worker which is common in the productivity (Syverson, 2011; Brandt *et al.*, 2012) and temperature-productivity literature (Hsiang, 2010; Dell *et al.*, 2012). Firms report value added directly in the data and it equals total production (including both sales and inventory) of all goods produced in the year valued at their market prices less the cost of all intermediate inputs employed in producing them. Using aggregate measures of productivity requires that prices do not reflect market power in either the primary or upstream input markets. We cannot guarantee this; however, nearby-city pollution is independent of firm-level market power in the focal city allowing us to consistently estimate pollution's effect on productivity via instrumented pollution. The mix of products is also not discernible from firm-level value added and may be correlated with local pollution levels. However, our instrumenting strategy also addresses this issue: nearby-city pollution is uncorrelated with the product-mix decisions of a firm in the focal city thereby removing any bias in the instrumented results.

As explained below, we impose a maximum distance of 1,800 kilometers in estimating the spillover decay function and 300 kilometers in the causal estimates of productivity effects. After merging the productivity, API, and weather data for the spillover estimates, the data include 60 focal cities that represent 26% of China's population. The total annual output of these cities is CNY 2.02 trillion (11.7% of

[13] Available at www.wunderground.com.

[14] A 2007 exchange rate of 7.6 is used throughout the paper.

China's annual GDP and 29% of China's manufacturing sector).[15] For the casual estimates, the data includes 88,716 firms in 47 focal cities with total annual output of CNY 1.35 trillion (7.8% of China's annual GDP and 20% of China's manufacturing sector). Although the sample of cities is not comprehensive these are major cities representing a significant fraction of manufacturing output and population.

3. Estimation

3.1 Overview of estimation approach

As we show below, reduced-form estimation of spillover effects on productivity produces inefficient estimates. This will also not provide estimates of the local causal effects to compare with. To overcome these two issues, we rely on the fact that the reduced-form effect equals the intensity of treatment (the effect of nearby- on focal-city pollution) multiplied by the causal effect of focal-city pollution on focal-city productivity. We call the effect of nearby- on focal-city pollution the "pollution decay function" since we allow it to vary as a function of distance. Letting P^n represent nearby-city pollution, P^f focal-city pollution, and Y^f the focal-city outcome (in our case productivity):

$$\begin{gathered} spillover\ of\ P^n on\ Y^f = \\ (pollution\ decay\ function\colon effect\ of\ P^n\ on\ P^f) \times (causal\ effect\ of\ P^f\ on\ Y^f). \end{gathered} \tag{1}$$

This follows because the causal effect estimated via 2SLS using nearby-city pollution as an instrument is (Angrist and Pischke, 2015: 107):

$$causal\ effect\ of\ P^f\ on\ Y^f = \frac{(spillover\ of\ P^n\ on\ Y^f)}{(pollution\ decay\ function\colon effect\ of\ P^n\ on\ P^f)}. \tag{2}$$

We therefore proceed in two steps. In the first step we estimate the pollution decay function using daily data. We allow the effect to vary at different distances with controls for weather and seasonality. In the second step we employ the M2SLS method to estimate the causal effect of focal-city pollution on focal-city productivity using annual data, instrumenting daily focal-city pollution with daily nearby-city pollution conditional on wind direction. This step estimates the local average treatment effect of pollution on productivity. We then multiply the estimates for the spillover decay function obtained in the first step by the instrumental variable coefficient from the second step to yield the spillover effect of nearby-city pollution on focal-city productivity according to Equation (1). We bootstrap to compute standard errors that account for estimation error across both steps. The spillover

[15] China's average annual real GDP over the seven-year sample period is CNY 17.27 trillion. The manufacturing sector accounts for roughly 40% of China's GDP.

decay function is estimated at the city level because pollution is measured at that level while the causal effects of pollution on productivity are estimated at the firm level because productivity is measured and occurs at the firm level.

An additional advantage of separating these two steps is that the first step relating nearby- to focal-city pollution can involve very complicated relationships that depend on pollution, wind patterns, and weather in highly nonlinear ways while preserving the linear relationship necessary for instrumenting in the second step. The next subsection describes the first step of the approach (estimating the pollution decay function) and the following subsection the second step (estimating the causal effect).

3.2 Step one: estimating the pollution decay function

The pollution decay function isolates the physical transport of PM_{10} between nearby and focal cities. If wind direction is orthogonal to omitted factors that jointly affect both nearby- and focal-city pollution, relating the two during periods when wind blows toward the focal city identifies these spillovers. We offer evidence that wind direction is orthogonal to these omitted factors when we present the results. It is also necessary to isolate time periods in which the wind blows toward the focal city versus away. In the sample, wind direction changes by more than 90 degrees in absolute value (and therefore blows in the opposite direction) from day-to-day on more than 25% of days (Appendix C shows the full distribution of the change in wind direction across days). Averaging over a longer time period risks mingling periods in which the wind blows toward and away from the focal city. Thus, it is imperative to use daily data to isolate imported from local pollution.

We follow the concentric rings approach from the urban economics literature to estimate the pollution decay function.[16] This approach estimates the spillover between a location and each of several concentric rings radiating outward from that location. We use a piecewise linear regression to implement this, allowing the slope and intercept to differ for each of the concentric rings. We define rings at every 50 kilometers indexed by $b = 1,2,3,\ldots,B$ and identify all the nearby cities within each ring (if at least one exists) for each focal city. That is, all nearby cities within 0 to 50, 50 to 100, . . . , $(B-1)$*50 to B*50 kilometers. We expand B far enough to ensure the decay function has plateaued or hit zero ($B = 36$ or 1,800 kilometers).

Having identified these focal-nearby city pairs, we then estimate the impact of nearby city n's PM_{10} on focal city f's PM_{10} level on day d of month m in year t by

[16] The urban economics literature documents the spatial decay effects of agglomeration economies and knowledge spillovers (Rosenthal and Strange, 2003; Fu, 2007; Henderson, 2007; Arzhagi and Henderson, 2008; Rosenthal and Strange, 2008).

estimating the following equation conditional on the wind blowing from the nearby to the focal city:

$$P_{td}^{f} = \mathrm{I}_b\left[\lambda_{1b} + \lambda_{2b} abs\left[cos\left(\theta_{td}^{fn}\right)\right] P_{td}^{n}\right] + \lambda_3 W_{td}^{f} + \omega_f + \kappa_{rtm} + \varepsilon_{td}^{fn},$$

$$\forall f, n \in \mathcal{F}, n \neq f, \forall b = 1, \cdots B, \quad (3)$$

where $\mathcal{F}$ is the set of all cities in the data, P_{td}^{f} and P_{td}^{n} are the pollution levels of focal city f and nearby city n on day d of year t, and W_{td}^{f} are daily weather controls that affect pollution in the focal city. The indictor variable I_b is set to one for distance band b if nearby city n is within distance band b. λ_{1b} allows the intercept to vary for each distance band. λ_{2b} are the coefficients of interest and capture the average physical transport of nearby-city pollution to the focal city within each band. An observation in this regression is a focal-nearby city pair on a particular day. We form all possible pairings of focal and nearby city cities within 1,800 kilometers. Since each focal city may have more than one nearby city across or even within bands this is a stacked regression with potentially multiple observations per focal city.

We follow Schlenker and Walker (2016) in weighting nearby-city pollution by the absolute value of the cosine of the angle.[17] This angle $\left(\theta_{td}^{fn}\right)$ is the difference between the wind direction and the direction of the ray from the nearby to the focal city on day d of year t. For example, in Figure 1 where the focal city lies at an angle of 21° from the nearby city, if the wind is blowing at $-19°$ then $\theta_{td}^{fn} = -40°$ or if the wind is blowing at 43° then $\theta_{td}^{fn} = 22°$. We include a day in estimation as long as the wind blows within a 90° arc on either side of the ray connecting the nearby to the focal city. This is illustrated in the shaded area of Figure 1 for the example in which the focal city lies at an angle of 21° from the nearby city. In this example a day is included as long as $-69° < \theta_{td}^{fn} < 111°$. The pollution decay function is therefore identified from variation along two dimensions: distance between focal and nearby city and wind direction angle.

[Insert Figure 1]

W_{td}^{f} includes daily averages of relative humidity and wind speed, daily total precipitation, and temperature bins as described below. We include focal-city fixed

[17] We weight by the angle because more nearby-city pollution is imported the more directly wind blows toward the focal city. Using data for $-90° \leq \theta \leq 90°$ for the nearest nearby-city within 300 kilometers, the correlation between $cos(\theta)$ and residuals from regressing focal-city pollution on nearby-city pollution and focal-city weather is 0.046 significant at better than the 0.01% level. This means that if nearby-city pollution is increased by one $\mu g/m^3$ while θ is moved from 90° (perpendicular to the focal city) to 0° (directly toward the focal city), imported pollution increases by 0.046 $\mu g/m^3$ (21% of the total 0.216 $\mu g/m^3$ spillover at 300 kilometers shown in Appendix G).

effects (ω_f) to control for any time-persistent unobserved factors affecting the pollution drift to a focal city. Region-by-year-by-month fixed effects (κ_{rtm}) control for seasonal factors that affect pollution drift in a region such as wind patterns. We follow Zhang *et al*. (2018) in grouping the provinces into each of seven regions as described in Appendix D. The error term (ε_{td}^{fn}) captures any unobserved factors affecting drift between the focal-nearby city pair on day d of year t. We cluster standard errors at the focal-city level to allow for serial correlation across time within a focal city. This also allows for heteroscedasticity introduced by focal cities having different numbers of nearby cities.

3.3 Step two: estimating causal effect of pollution on productivity

In the second step we estimate the causal effect of focal-city pollution on focal-city productivity. In the short run, high air pollution concentrations can lead to decreased lung function, irregular heartbeat, increased respiratory problems, nonfatal heart attacks, and angina.[18] Long-run cumulative exposure may lead to cardiopulmonary diseases, respiratory infections, lung cancer (EPA, 2004), and asthma (Neidell, 2004) that can surface in the short run. All of these health conditions may decrease physical stamina and lead to missed work days. Workers may also be absent from work to care for the young and elderly affected by pollution (Chay and Greenstone, 2003; Hanna and Oliva, 2015; Deryugina *et al*., 2019; Aragón *et al*., 2017). Increased mortality (Chen *et al*., 2013; Ebenstein *et al*., 2017) can lead to experienced workers being replaced by less experienced ones. Air pollution can also have psychological effects including lowering cognitive ability, altering emotions, and increasing anxiety (Levinson, 2012; Lavy *et al*., 2014; Pun *et al*., 2016; Chen *et al*., 2018) which would affect both physical and mental performance. While the estimates are unable to distinguish between these various channels they capture the effect of all of them.

3.3.1 Step two: identification

We focus here on identification issues related to productivity but the identification arguments apply to endogeneity issues that arise from outcomes more broadly. OLS estimates are subject to simultaneity and omitted variable biases. Even without any effect of pollution on productivity, cities with more output will produce more pollution. If pollution does lower productivity, the lower productivity will result in less pollution. Firms may also respond to the lowered labor productivity by substituting from labor to alternative inputs.

Omitted-variable biases due to local, time-varying conditions are also possible (firm fixed effects absorb any time-invariant effects). For example, high-productivity firms may implement advanced, lower-polluting technologies over time while low-

[18] See the EPA website: https://www.epa.gov/pm-pollution.

productivity firms do not. Spatial sorting could introduce spurious correlations. Firms may choose to enter in or relocate to cities with less severe pollution because it will raise their productivity or in cities with more severe pollution because they have lax environmental regulations and impose fewer costs (Becker and Henderson, 2000; Greenstone, 2002; Brunnermeier and Levinson, 2004). Governments may force firms to relocate and pollution inflow from other cities may affect these decisions (for example, moving firms away from areas that are typically upstream of densely-populated areas). Firm exit may be endogenous due to the reduced productivity that pollution brings. Workers may also systematically sort across cities. High-skilled workers generally have a higher willingness-to-pay for clean air which would lead to low-skilled workers being located disproportionately in dirtier cities (Chen *et al*., 2017; Lin, 2017). The inclusion of firm fixed effects means that only migrations of firms or workers during the sample period will bias the results.

We address these issues using nearby-city pollution that drifts to the focal city as an instrumental variable to identify the causal effect of local pollution on local productivity. To ensure exogeneity, we condition on the wind blowing from the nearby to the focal city.[19] Exogeneity also requires that wind direction timing is random with respect to nearby-city air pollution, conditional on controls, which we confirm below.

The inclusion restriction requires that the nearby city is close enough that significant amounts of pollution can drift from it to the focal city. To ensure this, we include only focal cities that have a nearby city sufficiently close. We consider maximum distance cutoffs ranging from 150 to 300 kilometers (our pollution decay function estimates confirm significant transport at these distances) and find robust results. There is a tradeoff in increasing the distance: it increases the available data but weakens the instrument's power. To also increase the instrument's power we include only the nearest nearby city for each focal city. As a result, even with a maximum distance of 300 kilometers the average distance between focal and nearby cities is only 106.5 kilometers.

The exogeneity condition requires that unobserved determinants of focal-city productivity are uncorrelated with the nearby city's pollution. This requires high-frequency data for two reasons. First, periods in which the wind imports pollution from outside must be isolated from those when it does not. To ensure this, in the instrumenting equation we condition on the wind blowing from the nearby to the focal city on a particular day. We offer evidence when we present the results that daily data succeeds in isolating periods when wind blows toward the focal city.

[19] When the wind blows toward the nearby city its pollution is not exogenous because greater focal-city output increases the nearby city's air pollution.

Conveniently, this high-frequency instrument is already available as it is required to estimate the pollution decay function.

Second, high-frequency data is required to ensure that common shocks do not affect both focal- and nearby-city output. Positive regional shocks to productivity could raise both cities' output thereby increasing nearby-city pollution as well. Alternatively, if focal- and nearby-city production are substitutes in output markets then output growth in a focal city will reduce nearby-city output and pollution. While common regional shocks are likely to induce correlated actions across cities over a long time period, they are unlikely to do so over a short time frame due to lags in shock propagation and delays in responses to those shocks. With the use of daily data, violating the exogeneity condition would require that shocks affect focal- and nearby-city productivity on a *daily* basis.

This addresses each of the potential endogeneity biases. Nearby-city pollution is uncorrelated with focal-city output in the absence of common regional shocks that are propagated and responded to on a daily basis. Trends in pollution and productivity would need to be correlated across the focal and nearby city on a daily basis to bias the estimates. Substitution away from labor and toward other inputs in response to imported pollution would need to occur on a daily basis. Similarly, firm entry, exit, or relocations and worker migrations in response to imported pollution would need to occur on a daily basis.[20]

This instrumenting strategy can be implemented using either M2SLS with daily data in the first stage or Wald 2SLS with annual averages in the first stage (in either case conditioning on wind direction). Appendix E shows formally that either approach produces unbiased estimates in the presence of a common shock to focal- and nearby-city output as long as it is of lower than daily frequency. However, there are two important differences between the two estimation approaches. M2SLS produces unbiased estimates in the first stage because intra-year common regional shocks to pollution (as opposed to output) can be controlled for using fixed effects while Wald 2SLS may produce biased estimates.[21] Second, M2SLS produces more efficient second-stage estimates as we demonstrate below. These two differences are also shown formally in Appendix E.

---

[20] For example, suppose a factory moved from a focal city to a nearby city mid-year. For the first half-year, the local pollution it produces would lower productivity but this would not affect our estimates since this pollution is uncorrelated with nearby-city pollution conditional on wind direction. In the second half-year, this would increase the pollution that drifts to the focal city from the nearby city. It would also decrease productivity in the focal city in the last half-year due to spillovers. Our estimates would capture this since we condition on wind direction.

[21] For M2SLS, these are controlled for by region-by-year-by-month fixed effects in the first stage. For Wald 2SLS the first stage is biased by these effects; however, the second stage remains unbiased because the predicted values from the first stage are uncorrelated with the common shocks to output that may be present in the second stage.

In the results we assess the effects of aggregating the instrument to lower and lower frequencies. Consistent with theoretical predictions in Appendix E, the first-stage coefficient becomes increasingly biased at lower frequencies due to common shocks to focal- and nearby-city pollution and the second-stage coefficient is less and less precise.

3.3.2 Step two: implementation

The outcome that we wish to estimate (productivity) is measured annually while the pollution instrument is daily. A standard way of proceeding is to estimate Wald 2SLS using annualized values (conditional on wind direction in the first stage). We show below that these estimates are very inefficient. Instead, we employ M2SLS which provides estimates that are consistent and asymptotically normal (Dhrymes and Lleras-Muney, 2006) provided that the groupings are independent of the structural error as they are when the grouping is a primitive (in our case grouping daily observations into years).[22] Theoretically, M2SLS can be more or less efficient but we show in our setting that it is much more efficient.

The first-stage equation predicts air pollution for firm i located in focal city f of region r on day d in month m of year t conditional on the wind blowing from the nearby to the focal city. While the spillover equation in step one uses city data, this equation uses firm data to be consistent with the firm data used in the second stage:

$$P_{itd}^{f} = \gamma_1 abs\left[cos\left(\theta_{itd}^{fN^*}\right)\right] P_{itd}^{N^*} + \gamma_2 W_{itd}^{f} + \alpha_i + \kappa_{rtm} + \epsilon_{itd}^{f}, \quad (4)$$

where P_{itd}^{f} is the pollution in firm i's focal city f on day d of year t, $\theta_{itd}^{fN^*}$ is the wind direction relative to the ray from the nearest nearby city to firm i's focal city on day d of year t, and $P_{itd}^{N^*}$ is the pollution level on that same day in focal city f's nearest nearby city $N^* \in \mathcal{F}$ within a maximum radius distance. If no nearby city is available for a focal city it is dropped from the estimation. Every nearby city is also a focal city although it might be paired with a different nearby city that is closer. We test the sensitivity of the results to maximum distance cutoffs ranging from 150 to 300 kilometers.[23] W_{itd} is a vector of daily weather variables faced by firm i on day d of year t. We include linear and quadratic functions of daily relative humidity, wind speed, and cumulative precipitation. We allow for a flexible, nonlinear function of temperature following Deschênes and Greenstone (2011) and Zhang *et al.* (2018) since it has been found to affect productivity (Zhang *et al.*, 2018). We construct

[22] Lleras-Muney (2005) applies M2SLS to estimate the causal impact of education on health, Massa and Žaldokas (2014) to estimate international demand for US bonds, and Jordan (2016) to estimate local environmental preferences on mine closures.

[23] Distances below 150 kilometers yielded insufficient data and distances above 300 kilometers yielded a weak instrument as we demonstrate below.

indicator variables for the daily average temperature below 0°, 5° intervals from 0 to 30°, and above 30° Celsius.

In defining whether the wind blows toward the focal city, we impose more stringent criteria than in the pollution decay function estimation to ensure a sufficient quantity of pollution is imported from the nearby city. This is necessary for the instrument to be powerful.[24] For the baseline estimates, we include a day if the wind passes within a 45° arc on either side of the ray connecting the two cities. We refer to this as the "middle" funnel. Figure 2 illustrates this for the example in which the focal city lies at an angle of 21° from the nearby city. In this case a day is included as long as $-24° < \theta_{td}^{fn} < 66°$ (the shaded region of the figure). We check the robustness of the results to arcs of $\pm 40°$ ("narrow" funnel) and $\pm 50°$ ("broad" funnel). As in the pollution decay function estimation, the nearby-city's pollution is weighted by the absolute value of the cosine of the angle.

[Insert Figure 2 here]

Firm fixed effects (α_i) capture time-persistent unobservables that affect firm i's pollution exposure. Since no firms switch focal cities or industries over the sample period, these also absorb city-specific and industry-specific time-invariant factors that affect local pollution. Region-by-year-by-month fixed effects (κ_{rtm}) control for any year-month specific unobservables affecting the pollution in a region. We cluster standard errors at the focal-city level to allow for spatial correlation for all firms within each focal city and serial correlation across days within a focal city over time.

This equation differs from the pollution decay function (Equation (3)) in two ways. First, in order to ensure the power of the instrument, Equation (4) restricts estimation to shorter distances (a maximum of 300 kilometers), it utilizes only the nearest nearby city, and includes only days when the wind direction is within a funnel rather than within a half-circle. This maximizes the potential for the nearby city's pollution to drift to and affect the focal city. The objective of Equation (3) is to estimate spatial decay and it therefore utilizes all of the nearby cities to a focal city, utilizes all days of wind direction within a half-circle, and extends the measurement of these spillovers to a much greater distance. Second, Equation (3) also allows for a much more flexible functional form for estimating the spillover decay function than the linear restriction that 2SLS imposes on Equation (4).

Using the results from estimating Equation (4), we compute predicted values $\hat{P}_{itd}^{f}$ for each day included in the estimation (wind blowing toward the focal city) and

[24] Footnote 18 provides evidence that nearby-city pollution is a stronger instrument when the wind blows more directly in the direction of the focal city.

average them over days within each firm-year to obtain instrumented pollution for the second-stage: $\hat{\bar{P}}_{it}^{f}$. The second-stage equation is:

$$\ln\left(Y_{it}^{f}\right) = \beta_1 \hat{\bar{P}}_{it}^{f} + \gamma_2 \bar{W}_{it}^{f} + \alpha_i + \delta_{rt} + \eta_{it}^{f}, \quad (5)$$

where Y_{it}^{f} is value added per employee for firm i in the focal city f in year t and $\bar{W}_{it}^{f}$ contains the weather controls from the first stage averaged over all days within each firm-year (i.e., averages of the linear and quadratic functions of non-temperature variables and temperature bins containing the fraction of days in which the average temperature is below 0°, in 5° intervals from 0 to 30°, and above 30° Celsius).[25]

Firm fixed effects α_i capture time-persistent firm attributes that affect labor productivity. Region-by-year fixed effects (δ_{rt}) capture time-varying, regional shocks to firm output. The error term (η_{it}) includes time-varying, firm-level shocks to productivity. We cluster standard errors at the focal-city level to allow for serial correlation within each firm over time and spatial correlation within each city. We adjust for the error introduced in the first-stage estimation using a block bootstrap as in Schlenker and Walker (2016) with 100 iterations.

4. Results

Before we show the results of our approach we establish that a straightforward reduced-form regression of focal-city productivity on nearby-city pollution produces inefficient estimates. To do so, we aggregate the nearby-city pollution to the annual level conditional on wind direction, weighted by the cosine of the wind-direction angle, and including control variables corresponding to those in the M2SLS procedure.[26] Appendix F graphs the results converting them to their monetary impact. It shows the effects of a one $\mu g/m^3$ annual increase in nearby-city PM_{10} within a distance band (holding all others constant) on the average firm's annual productivity along with the 95% confidence interval in red, dashed lines. All the effects except for the 0-50 kilometer distance band are close to zero and almost all are insignificant. Given this lack of precision, we now turn to our approach.

---

[25] To ensure the exclusion restriction is met, the first-stage equation must include the non-averaged values of all the exogenous variables from the second stage. The weather controls in the second stage ($\bar{W}_{it}^{f}$) are yearly averages of the linear and quadratic terms of all non-temperature variables in the first stage. For the temperature variable, the bins in the second stage are annual averages of the daily indicator variables included in the first stage. The firm fixed effects remain the same as in the first stage. Finally, the region-by-year fixed effects included in the second stage are averages of the region-by-year-by-month fixed effects in the first stage.

[26] An alternative reduced-form approach would be to regress annual productivity on daily nearby-city pollution but this would involve over two billion observations in order to estimate as a nonlinear function of distance.

We report the first-step estimates (pollution decay function) followed by the second-step estimates (causal effects of focal-city air pollution on focal-city productivity) and then combine the results from these two steps to calculate the spillover effects of nearby-city pollution on focal-city productivity. After this, we demonstrate the advantage of the M2SLS procedure. In particular, we show that estimating causal effects using Wald 2SLS with annual data produces insignificant second-stage results and biased first-stage results. We offer supporting evidence that this is due to aggregating the high-frequency data to a lower frequency.

4.1 Pollution decay function

To estimate the pollution decay function we include all focal cities with at least one nearby city within 1,800 kilometers. This distance was chosen because it was far enough that the spillover effects were indistinguishable from zero.[27] We use all cities that have daily API and weather data available from 2001 to 2007. This yields 60 unique cities in a panel which is unbalanced because API data was not reported for some cities in the earlier years. There are some days with missing API or wind data but these are limited (all cities have at least 335 days of data in each year) and we believe are due to random non-reporting.

Table 1 shows the summary statistics for the pollution spillover data. There are 2,586 focal-nearby-city pairs (about 43 nearby cities for each focal city). If city B is a focal city for A then A is also a focal city for B. The focal cities' PM_{10} levels average 97.5 and exhibit significant variation. Wind blows toward the focal city on 52.1% of the days and PM_{10} is the dominant pollutant on 92% of the days for the focal cities. The mean distance between cities (1,004 kilometers) is about one-half the maximum allowed distance.

[Insert Table 1 here]

The solid, black line in Appendix G shows the λ_{2b} coefficients from estimating Equation (3) along with the 95% confidence interval in red, dashed lines. These are the effects of a one μg/m^3 increase in PM_{10} in nearby cities conditional on wind blowing directly toward the focal city $\left(\theta_{td}^{fn}=0\right)$. The effect in each distance band is conditional on holding PM_{10} in other bands constant. Roughly 45% of pollution drifts from nearby cities that are within 50 kilometers and more than 18% at 400 kilometers.

The solid, black line in Figure 3 plots the effect of a one μg/m^3 *annual* increase in nearby-city PM_{10} along with the 95% confidence interval in red, dashed lines (for

[27] Re-estimating with a maximum radius of 1,200 kilometers (just above the point at which the effects hit zero) yields almost identical coefficients and standard errors.

clarity we plot only to a distance of 1,200 kilometers). This adjusts the coefficients using the empirical distribution of θ_{td}^{fn}. That is, for the fact that the wind blows toward the average focal city on only 52.1% of days in a year and does not always blow directly towards the focal city. Again, this is the effect of increasing PM_{10} in the distance band conditional on holding pollution constant in all other bands.[28] The spillover effect within 50 kilometers is 0.355. That is, a one $\mu g/m^3$ annual increase in PM_{10} in all nearby cities within 50 kilometers, but not in any other distance band, increases annual focal city pollution by 0.355 $\mu g/m^3$. Similarly, a one $\mu g/m^3$ annual increase in PM_{10} in all nearby cities within 50 to 100 kilometers, but not in any other band, increases annual focal city pollution by 0.185 $\mu g/m^3$. A similar analysis applies to all the further distance bands. These effects are for the average focal city in the sample given average weather. Spillovers drop somewhat quickly and smoothly from 0.355 at 50 kilometers to 0.062 at 600 kilometers after which they fall more slowly to zero at about 1,000 kilometers.

[Insert Figure 3 here]

4.2 Randomness of daily wind data

Before estimating the causal effect of pollution on productivity, we check the randomness of wind direction with respect to pollution. To ensure that the instrument is exogenous we must exclude days in which the wind does not blow from the nearby to the focal city. If wind direction is not randomly distributed with respect to the distribution of nearby-city air quality, conditional on control variables, this may bias the coefficients.[29] Appendix H compares cumulative distribution functions (cdfs) of nearby-city air pollution conditional on the control variables used in the first stage of the M2SLS procedure for all days versus excluded days using the 150-, 200-, 250-, and 300-kilometer distance cutoffs in choosing nearby cities. The cdfs are very similar for all cutoffs.[30]

---

[28] It would be useful to compare the local effect to spillovers from raising pollution in all nearby cities simultaneously. However, to do so using our estimates requires making arbitrary assumptions about the degree to which pollution from a nearby city affects other nearby cities that are between it and the focal city. Alternatively, one could estimate spillovers including interaction effects between each distance band and all closer distance bands to estimate these "pass-through" effects. However, the number of independent variables required makes this infeasible with more than a few distance bands.
[29] This highlights the importance of the control variables. For example, in northern regions of China air quality is worse in the winter than in other seasons. If wind directions are systematically different in winter than other times of the year this will introduce bias in the absence of control variables. In this example, the region-by-year-by-month fixed effects capture this region-specific seasonality.
[30] A two-sample Kolmogorov-Smirnov test rejects the null hypothesis of the equality of distributions for three of the radius cutoffs; however, the magnitude of the differences is very small. For the 200-kilometer radius the difference is significant at the 1.8% level but the maximum difference is only 0.016. For the 250-kilometer radius the difference is significant at the 3.0% level but the maximum difference is only 0.014 and for the 300-kilometer radius the difference is significant at the 3.9% level

4.3 Effect of local air pollution on local labor productivity

In this subsection we estimate the causal effect of focal-city pollution on focal-city labor productivity using nearby-city pollution as an instrument. In choosing which nearby cities to include, we check robustness to maximum distances from the focal city of 150, 200, 250, and 300 kilometers. There is a tradeoff as this distance increases. There are more data available to identify the effects thereby increasing their precision; however, the instrument is weaker because nearby-city pollution has less effect on focal-city pollution. Below 150 kilometers there were insufficient data to identify effects and we show that beyond a distance of 300 kilometers the instrument is no longer powerful. Unlike the spillover estimates, we choose the nearest nearby city to the focal city, if one exists, within the maximum distance to maximize the instrument's power.

Table 2 shows summary statistics for the main variables for the 150- and 300-kilometer radiuses. The top panel summarizes the first-stage data which are at the firm-day level. The summary statistics are fairly similar across the two distance cutoffs. The PM_{10} levels are high enough to potentially affect productivity. The annual mean is 112 $\mu g/m^3$ compared to a World Health Organization (WHO) recommended guideline of 20 $\mu g/m^3$ annual average and many days exceed the WHO guideline of 25 $\mu g/m^3$ for a 24-hour average (World Health Organization, 2006). As the cutoff increases from 150 to 300 kilometers, the number of focal cities increases from 30 to 47. The average distance between nearby and focal cities does not increase much because we use the nearest nearby city for each focal city. The bottom panel summarizes the second-stage data which are at the firm-year level. The data exhibit significant variation in value-added per employee. Appendix I shows summary statistics for the 200- and 250-kilometer radiuses which are similar.

[Insert Table 2 here]

Panel A of Table 3 shows OLS results that do not address the endogeneity of air pollution. The firm-year data included here correspond to those included in the second stage of M2SLS estimation described below. For all four distance cutoffs, the coefficients on PM_{10} are insignificantly different from zero and for all but the 150-kilometer the point estimates themselves are close to zero.

We now turn to M2SLS estimates. Panel B shows the results of estimating the first-stage equation (Equation (4)) using PM_{10} of the focal city's nearest nearby city as an instrument conditional on wind blowing toward the focal city within the middle

but the maximum difference is only 0.013. For a 150-kilometer radius the difference is not quite significant (10.8%) but the maximum difference is only 0.014. This is an example of Simpson's Paradox in which a large amount of data (for the 200-kilometer radius there are 55,088 observations) results in statistical significance for even small differences.

funnel. This estimation is at the firm-day level and the wind is within the middle funnel on about one-fourth of the days. The results reveal a strong instrument. A one μg/m³ increase in a nearby city's PM_{10} increases the focal city's PM_{10} by between 0.70 and 0.72 with a high level of significance.[31] This is not too far from the theoretical upper bound of 1.0 because it uses only the nearest nearby city and pertains to days when the wind is blowing directly toward the focal city $(\theta_{td}^{fN*} = 0)$. The physical transport of pollution is lower when the wind is not blowing directly toward the focal city or from nearby cities that are further away. The Kleibergen-Paap Wald rk (KP) *F*-statistic (Kleibergen and Paap, 2006) for weak identification significantly exceeds the Stock-Yogo critical value of 16.38 for all four cutoffs.[32]

Panel C shows the second-stage estimates of Equation (5) at the firm-year level using the average values of the predicted pollution from the first stage as an instrument and controlling for weather and region-by-year fixed effects. The estimated coefficients of PM_{10} are negative and significant for all but the 150-kilomoter cutoff. The estimates become more significant as the cutoff increases consistent with more data used in estimation. The coefficients are fairly consistent across the four cutoffs and imply that a one μg/m³ annual increase in PM_{10} decreases productivity by 0.26 to 0.34%. Evaluated at the mean focal-city PM_{10} in each subsample, these estimates imply elasticities of labor productivity with respect to air pollution of -0.29 to -0.35.[33]

These results are consistent with the instrument attenuating an upward endogeneity bias. The results also imply that improving air quality generates substantial productivity benefits. Using the 300-kilometer cutoff data and estimates, a 1% reduction in PM_{10} increases per-firm productivity for the average firm by CNY 47,700 (USD 6,276) annually. Throughout the remainder of the paper we use the 300-kilometer estimate as our preferred since it is the most significant and is close to the midpoint of the estimates from the three significant cutoffs.

[Insert Table 3 here]

Column 2 of Appendix J reports results of a counterfactual test of the instrument. It uses M2SLS with the middle funnel and a 300-kilometer radius but conditions on wind blowing away from the focal city in instrumenting for focal-city PM_{10}. The first stage results (shown in Panel A) are nearly identical to those using the baseline model (reproduced in Column 1). This is not surprising: focal-city pollution should

[31] These coefficients exceed the estimates even at 50 kilometers in the spillover decay function (0.45 from Appendix G) because here we estimate using a funnel that is twice as narrow.
[32] Stock and Yogo (2005) critical values apply when model errors are independent and identically distributed. No critical values are available for the case when the model allows for standard errors that are robust to heteroskedasticity and clustering.
[33] Mean annual PM_{10} (unconditional on wind direction) in the second-stage data is 104.1 for 150-, 111.3 for 200-, 103.1 for 250-, and 104.1 for 300-kilometer radius.

have the same effect on nearby-city pollution when wind is blowing toward the nearby city as vice-versa. The second-stage results (shown in Panel B) are very different than the baseline results. The coefficient is much lower in magnitude and insignificant consistent with the instrument not addressing endogeneity bias. In fact, the estimates are similar to the OLS results in Panel A of Table 3.

Appendix J contains other robustness checks of the estimates using the 300-kilometer cutoff. Column 3 uses a narrow funnel (an 80° arc). The point estimate is slightly smaller and is significant only at the 16% level due to the loss of data in the first stage. Employing a broad funnel (a 100° arc) with more data in Column 4 produces a somewhat more significant and larger effect than the baseline estimate. Dropping days with API below 50, for which the major pollutant is unknown, lowers the coefficient somewhat (Column 5). This is presumably due to years with a relatively high number of low-pollution days corresponding to years with a relatively high proportion of high-productivity days. Column 6 shows the importance of including weather controls. Without them, the coefficient is lower and no longer significant either because of their effect on the instrumented values or as a control for factors affecting productivity. Including the potentially manipulated range of API (Column 7) produces almost identical results to the baseline. Including year-by-month rather than region-by-year-by-month fixed effects in the first stage (Column 8) yields similar results to the baseline but even more significant while including region-by-year fixed effects in the first stage results in somewhat different estimates with less significance (Column 9).[34] Therefore, the estimates are sensitive to controlling for overall seasonality more so than region-specific effects.

Appendix K provides supporting evidence for the choice of 300 kilometers as the maximum distance for the nearest nearby city to include as an instrument. Column 1 reproduces the baseline estimates. Column 2 estimates M2SLS using as an instrument pollution in the nearest nearby city for each focal city that is further than 300 but less than 350 kilometers away and using the middle funnel in defining whether wind blows toward the focal city. Columns 3 through 5 expand the data by increasing the range of distances for the nearest nearby cities. The average distance between the focal and nearby cities increases from 106.5 kilometers in the baseline estimates compared to more than 323.9 kilometers in the counterfactual estimates. The first-stage results in Panel A reflect the reduced power of the instrument compared to the baseline. The coefficient is about half that in the baseline estimates and the KP *F*-statistic is much lower. The second-stage coefficients (Panel B) are all insignificant consistent with a weak instrument.

[34] We experimented with using province-by-year-by-month fixed effects but the model was too saturated. There is an average of only 1.5 cities per province in the data.

4.4 Spillover effect of nearby-city pollution on focal-city labor productivity

As shown in Section 2.1, multiplying the first-step spillover decay function by the second-step causal effects yields the spillover effects of nearby-city pollution on focal-city productivity. To obtain appropriate standard errors clustered at the city level for these spillover effects we employ a block bootstrap with 100 iterations.[35] We estimate this using a 300-kilometer cutoff and middle funnel for the instrument in the M2SLS estimation.

Figure 4 summarizes the results converting them to the monetary impact for the average firm's annual productivity on an average weather day. The solid, black line shows the effect of a one μg/m^3 annual increase in nearby-city PM_{10} in that distance band (holding pollution in all other bands constant) on focal-city productivity with 95% confidence intervals shown in dashed, red lines. Since these are annual productivity effects this assumes a one μg/m^3 increase in nearby-city PM_{10} for the entire year and adjusts for the empirical distribution of wind direction during the year. The costs are CNY 16,248 (USD 2,138) for nearby cities within 50 kilometers and decline fairly quickly and smoothly to CNY 2,847 (USD 375) for nearby cities at 550 to 600 kilometers. Beyond this, the spillovers decline slowly to approach zero at about 1,150 kilometers (for clarity we plot only to 1,200 kilometers). In comparison the effect of local sources of PM_{10} on productivity is CNY 45,809 (USD 6,028).

[Insert Figure 4 here]

While the spillover decay function estimates alone tell us the relative tradeoff between local and extra-territorial effects, they do not tell us the absolute amounts at stake. This requires both steps of the procedure. For example, if PM_{10} increases by one μg/m^3 annually in both a focal city and a nearby city located at 90 kilometers, productivity falls by CNY 45,809 annually for the average firm due to local sources of pollution and another CNY 8,494 due to imported pollution. The latter is smaller because pollution dissipates as it drifts and the wind blows directly toward the focal city only part of the time. These absolute costs can be used to determine the geographic scope of environmental regulation necessary to internalize externalities that are above a given cost.

These results can also be used to calculate Coasian prices. Consider Tianjin which is located 107 kilometers from Beijing and let θ_{td}^{BT} be the angle of the wind relative to the ray from Tianjin to Beijing. If each city were assigned rights to keep its city free of other cities' air pollution, Tianjin would have to compensate Beijing CNY

[35] Specifically, for each iteration we draw (with replacement) a block bootstrap by city. In the first step (spillover decay function) we use all days in all years for these cities. In the second step (causal effects) we use all firms and all days in all years for the sampled cities.

35*$abs[cos(\theta_{td}^{BT})]$ times the number of firms in Beijing on each day when $-90^{\circ} \leq \theta_{td}^{BT} \leq 90^{\circ}$, where 35 is the λ_{2b} coefficient from Equation (3) multiplied by the annual causal effect converted to a daily cost.[36] Similarly, on days when the wind blows toward Tianjin, Beijing would have to compensate Tianjin 35*$abs[cos(\theta_{td}^{TB})]$ times the number of firms in Tianjin for each μg/m^3 of PM_{10} that Beijing produces on a day when the wind blows between $-90^{\circ} \leq \theta_{td}^{TB} \leq 90^{\circ}$where θ_{td}^{TB} is the angle of the wind relative to the ray from Beijing to Tianjin. Some of the pollution blowing from Beijing to Tianjin may have originated in other cities before being passed on to Tianjin. These other cities would compensate Beijing using the same approach so that Beijing's net payment would correspond only to the pollution that it originated.

4.5 Wald 2SLS estimates

An alternative to the M2SLS procedure is to combine the first-step estimates of the pollution decay function using daily data with causal estimates based on Wald 2SLS. Estimating Wald 2SLS requires aggregating the first-stage data to match the annual data used in the second stage. We aggregate the first-stage data by taking firm-year averages conditional on wind blowing toward the focal city (i.e., computing mean values of focal-city pollution and cosine-weighted nearby-city pollution using only days when the wind blows toward the focal city). We also include weather controls, firm and region-by-year fixed effects, and cluster standard errors by focal city to be consistent with the M2SLS estimates. Table 4 shows the results at the different distance cutoffs using the middle funnel.

The coefficients for the first-stage results (Panel A) are all significant but are opposite of the expected sign. This is because when there is variation within groups, grouped estimation identifies parameters that differ from those in ungrouped estimation (Angrist and Pischke, 2011: 314). Appendix L shows scatter plots that relate focal-city PM_{10} conditional on first-stage control variables and nearby-city PM_{10} for daily values versus annual average values along with fitted regression lines. In both cases we condition on wind blowing toward the focal city. The daily plot shows a clear positive relationship between the city pairs' pollution values. The primary effect of aggregating to the annual level is a loss of precision in the relationship but the relationship also becomes negative. This results from common shocks to focal- and nearby-city pollution that are negatively correlated and occur at lower frequencies than daily. As a result, the first-stage coefficient is biased downward (see Equation

[36] The λ_{2b} coefficient is 0.279 for nearby cities between 100 and 150 kilometers away. The annual causal effect is CNY 45,809 or CNY 125 daily. Multiplying these two numbers yields CNY 35.

(A10) of Appendix E for a formal exposition). The results also suggest weak instruments with all of the KP *F*-statistics below the critical value of 16.38.[37]

As Appendix E shows, Wald 2SLS still produces unbiased second-stage estimates even with a biased first-stage coefficient. The first-stage fitted values from either Wald 2SLS or M2SLS reflect the component of focal-city pollution that is due to variation in nearby-city pollution. However, Wald 2SLS may be less efficient. M2SLS is more efficient because it uses disaggregated data in the first stage thereby utilizing more information; however, the grouping of the first-stage predicted values changes the nature of the first stage errors and their relationship to the second-stage errors which could decrease efficiency (Dhrymes and Lleras-Muney, 2006). Appendix E provides the formal statistical test of whether M2SLS is more efficient in our setting based on Dhrymes and Lleras-Muney (2006). The test statistic is 1,735.1 compared to a cutoff value of 1.64 for a 5% level of significance indicating M2SLS estimates are much more inefficient than Wald 2SLS.

[Insert Table 4 here]

M2SLS is much more efficient in our setting because vastly more information is used in the first stage of M2SLS than in Wald 2SLS. This overwhelms any loss of efficiency due to correlations between the annualized first-stage and second-stage residuals. The same gain in efficiency is likely to be achieved when applying M2SLS to other outcomes because of the much greater information in daily data (using the middle funnel one-fourth of days are used implying 91 times as many observations with daily than annual data). Consistent with the lower efficiency of Wald 2SLS, the second-stage coefficients in Panel B of Table 4 are insignificant for all four cutoffs. We now investigate this loss of efficiency further.

Table 5 shows how the level of aggregation in the first-stage affects the estimates of the causal effects of pollution on productivity (second-stage estimates are all at the firm-year level). These estimates use the 300-kilometer cutoff in choosing the nearest-nearby city, apply the middle funnel in choosing which days to include in the first-stage, and include the same controls as the baseline estimates except that region-year fixed effects are used rather than region-by-year-by-month.[38] Column 1 of the table uses firm-day data in the first stage conditional on wind blowing toward the focal city. This specification is the same as the baseline except that region-year fixed effects are used. As showed earlier, the causal effects are somewhat lower and

[37] Consistent with a single instrument that is very significant, a standard Cragg-Donald (1993) test overwhelmingly rejects the null hypothesis of weak instruments (e.g., a test statistic of 64,400 for the estimates using a 300-kilometer radius). However, the KP tests which adjusts for correlation in the errors results in a much lower test statistic.

[38] Region-by-year-by-month fixed effects are not used since they cannot be included once data is aggregated for periods longer than one month.

less significant using region-by-year fixed effects than in the baseline estimates using region-by-year-by-month fixed effects. This highlights another advantage of using daily data: finer controls in the first stage can lead to more efficient estimates.

Column 2 aggregates the first-stage data to the weekly level conditional on wind direction (i.e., averages all days when wind is blowing toward the focal city across each week). The first-stage coefficient remains similar and the second-stage coefficient is similar in magnitude but is significant only at the 11.0% level. Columns 3 through 6 aggregate in a similar way to the monthly, quarterly, semiannual, and annual levels (the last is the Wald estimates discussed above). Fairly clear patterns emerge as the level of aggregation is increased. The first-stage coefficient declines in magnitude (and turns negative with annual aggregation) while the second-stage coefficients become less and less significant. These results suggest that daily data is necessary to generate sufficient variation for precise estimates.

[Insert Table 5 here]

5. Conclusion

We provide a methodology for estimating the causal effect of air pollution spillovers on outcomes that are measured with lower frequency than pollution and weather data. Measuring air pollution spillovers requires high-frequency (such as daily) data to ensure that shifts in wind direction are properly captured, but outcome variables are often available on only an annual basis.

We proceed by estimating the pollution decay function at high frequency separately from the causal effects and estimating the causal effects using a mixed two-stage least squares (M2SLS) procedure using high-frequency changes in imported pollution from nearby cities as an instrument. The M2SLS procedure allows high-frequency data for the instrumenting in the first stage but low-frequency outcome data in the second stage. This estimation is a natural by-product of estimating the spillover decay function since this also requires high-frequency wind and pollution data. We show that typical Wald 2SLS fails in estimating causal effects due to the aggregation of pollution data over a long period and the resulting loss of efficiency.

Use of high-frequency data also allows spillovers to be examined at relatively short distances while minimizing the chance of spurious correlation from regional and seasonal shocks to the outcome variable. This allows an examination of spillovers between cities that are geographically close but administratively distinct and therefore potentially suffer from a free-rider problem in pollution production.

While we apply our procedure to quantify spillover effects of PM_{10} on productivity, our procedure can easily be adapted to estimate the spillover effects for other pollutants and on any outcome for which data is of a lower frequency than the pollution and weather data. For example, if only annual health measures are available the instrumenting technique works as long as daily pollution and weather data are available. It is also potentially applicable to estimating outcomes over periods longer than one year.

While previous papers document the presence of spillovers, our paper specifically quantifies how their intensity varies with distance— a necessary input for determining the scope of administrative control necessary to internalize externalities. PM_{10} spillovers in China are large and extend quite far suggesting the need to coordinate environmental policies at the supra-provincial level.

References

Altindag, D. T., D. Baek, and N. Mocan (2017). "Chinese Yellow Dust and Korean Infant Health," *Social Science & Medicine*, 186, 78 – 86.

Anderson, M. L. (2020). "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality," *Journal of the European Economic Association*, forthcoming.

Andrews, S. Q. (2008). "Inconsistencies in Air Quality Metrics: 'Blue Sky' Days and PM10 Concentrations in Beijing," *Environmental Research Letters*, 3, 1 – 14.

Angrist, J. D. and J. S. Pischke (2011). *Mostly Harmless Econometrics*. Princeton: Princeton University Press.

Angrist, J. D. and J. S. Pischke (2015). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton: Princeton University Press.

Aragón, F., J. J. Miranda, and P. Oliva. (2017). "Particulate Matter and Labor Supply: The Role of Caregiving and Non-Linearities," *Journal of Environmental Economics and Management*, 86, 295 – 309.

Arzaghi, M. and J. V. Henderson (2008). "Networking off Madison Avenue," *Review of Economic Studies*, 75, 1011 – 1038.

Banzhaf, H. S. and B. A. Chupp (2012). "Fiscal Federalism and Interjurisdictional Externalities: New Results and an Application to US Air Pollution," *Journal of Public Economics*, 96, 449 – 464.

Barwick, P.J., S. Li, D. Rao, and N.B. Zahur (2018). "The Morbidity Cost of Air Pollution: Evidence from Consumer Spending in China," *NBER Working Paper 24688*.

Becker, R. and J. Henderson. (2000). "Effects of Air Quality Regulations on Polluting Industries," *Journal of Political Economy*, 108, 379 – 421.

Bošković, B. (2015). "Air Pollution, Externalities, and Decentralized Environmental Regulation," working paper.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2012). "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 97, 339 – 351.

Brunnermeier, S. and A. Levinson. (2004). "Examining the Evidence on Environmental Regulations and Industry Location," *Journal of Environment & Development*, 13, 6 – 41.

Cai, H., Y. Chen, and Q. Gong (2016). "Polluting thy Neighbor: Unintended Consequences of China' s Pollution Reduction Mandates," *Journal of Environmental Economics and Management*, 76, 86-104.

Cai, H. and Q. Liu. (2009). "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms," *The Economic Journal*, 119, 764 – 795.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016). "Particulate Pollution and the Productivity of Pear Packers," *American Economic Journal: Economic Policy, 8,* 141 – 69.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2019). "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," *American Economic Journal: Applied Economics,* 11, 151 – 172.

Chay, K. Y. and M. Greenstone. (2003). "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *Quarterly Journal of Economics*, 118, 1121 – 1167.

Chen, S., P. Oliva, and P. Zhang (2017). "The Effect of Air Pollution on Migration: Evidence from China," *NBER Working Paper 24036*.

Chen, S., P. Oliva, and P. Zhang. (2018). "The Effect of Air Pollution on Mental Health: Evidence from China," *NBER Working Paper 24686.*

Chen, Y., A. Ebenstein, M. Greenstone, and H. Lie. (2013). "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy," *Proceedings of the National Academy of Sciences of the United States of America*, 110, 12936 – 12941.

Coase, R. H. (1960). "The Problem of Social Costs," *Journal of Law and Economics*, 3, 1 – 44.

Cragg, J. G. and S. G. Donald (1993) "Testing Identifiability and Specification in Instrumental Variable Models," *Econometric Theory*, 9, 222 – 240.

Dell, M., B. F. Jones, and B. A. Olken (2012). "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 4, 66 – 95.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif (2019). "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction," *American Economic Review*, 109, 4178 – 4219.

Deschênes, O. and M. Greenstone (2011). "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US," *American Economic Journal: Applied Economics*, 3, 152 – 185.

Dhrymes, P. J. and A. Lleras-Muney (2006). "Estimation of Models with Grouped and Ungrouped Data by Means of '2SLS'," *Journal of Econometrics*, 133, 1 – 29.

Ebenstein, A., Fan, M., Greenstone, M., He, G., and Zhou, M. (2017). "New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy," *Proceedings of the National Academy of Sciences*, *114*(39), 10384 – 10389.

Eichner, T. and M. Runkel (2012). "Interjurisdictional Spillovers, Decentralized Policymaking, and the Elasticity of Capital Supply," *The American Economic Review*, 102, 2349 – 2357.

EPA (U.S. Environmental Protection Agency). (2004). "Air Quality Criteria for Particulate Matter."

Federal Register (2017). "Revisions to the Guideline on Air Quality Models: Enhancements to the AERMOD Dispersion Modeling System and Incorporation of Approaches to Address Ozone and Fine Particulate Matter," 40 CFR Part 51, 5182 – 5235.

Fell, H. and D. T. Kaffine (2014). "Can Decentralized Planning Really Achieve First-Best in the Presence of Environmental Spillovers?" *Journal of Environmental Economics and Management*, 68, 46 – 53.

Freeman, R., W. Liang, R. Song, and C. Timmins (2019). "Willingness to Pay for Clean Air in China," *Journal of Environmental Economics and Management*, 94, 188 – 216.

Fu, S. (2007). "Smart Café Cities: Testing Human Capital Externalities in the Boston Metropolitan Area," *Journal of Urban Economics,* 61, 86 – 111.

Fu, S., V. B. Viard, and P. Zhang (2018). "Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China," SSRN Working Paper.

Ghanem, D. and J. Zhang (2014). "'Effortless Perfection:' Do Chinese Cities Manipulate Air Pollution Data?" *Journal of Environmental Economics and Management*, 68, 203 – 225.

Graff Zivin, J. and M. Neidell. (2012). "The Impact of Pollution on Worker Productivity," *American Economic Review*, 102, 3652 – 3673.

Greenstone, M. (2002). "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and Census of Manufactures," *Journal of Political Economy*, 110, 1175 – 219.

Hanna, R. and P. Oliva. (2015). "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico," *Journal of Public Economics*, 122, 68 – 79.

He, J., H. Liu, and A. Salvo (2019). "Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China," *American Economic Journal: Applied Economics*, 11, 173 – 201.

Henderson, J. V. (2007). "Understanding Knowledge Spillovers," *Regional Science and Urban Economics*, 37, 497 – 508.

Herrnstadt, E., A. Heyes, E. Muehlegger, and S. Saberian (2019). "Air Pollution and Criminal Activity: Evidence from Chicago Microdata," working paper.

Hsiang, S. M. (2010). "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America," *Proceedings of the National Academy of Sciences*, 107, 15367 – 15372.

Jia, R. (2017). "Pollution for Promotion," working paper.

Jia, R. and H. Ku (2019). "Is China's Pollution the Culprit for the Choking of South Korea? Evidence from the Asian Dust," *The Economic Journal*, 129, 3154 – 3188.

Jin, H., Y. Qian, and B. R. Weingast. (2005). "Regional Decentralization and Fiscal Incentives: Federalism, Chinese Style," *Journal of Public Economics*, 89, 1719 – 1742.

Jordan, B.W. (2016). "Can we Dig It? An Empirical Study of Environmental Attitudes and Social License to Operate in Mineral Extraction," working paper.

Kahn, M. E., P. Li, and D. Zhao (2015). "Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives," *American Economic Journal: Economic Policy*, 7, 223 – 242.

Kleibergen, F. and Paap, R. (2006). "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133, 97 – 126.

Lavy, V., A. Ebenstein, and S. Roth. (2014). "The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation," NBER Working Paper No. 20648.

Levinson, A. (2012). "Valuing Public Goods using Happiness Data: The Case of Air Quality," *Journal of Public Economics*, 96, 869 – 880.

Li, H. and L. Zhou (2005). "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China," *Journal of Public Economics*, 89, 1743 – 1762.

Lin, Y. (2017). "The Long Shadow of Industrial Pollution: Environmental Amenities and the Distribution of Skills," working paper.

Lipscomb, M. and A. M. Mobarak (2017). "Decentralization and Pollution Spillovers: Evidence from the Re-drawing of County Borders in Brazil," *The Review of Economic Studies*, 84, 464 – 502.

Lleras-Muney, A. (2005). "The Relationship between Education and Adult Mortality in the U.S.," *Review of Economic Studies*, 72, 189 – 221.

Massa, M. and A. Zaldokas (2014). "Investor Base and Corporate Borrowing: Evidence from International Bonds," *Journal of International Economics*, 92, 95 – 110.

Ministry of Civil Affairs (2017). "The Brochure of Administrative Divisions of Townships in China."

Moussiopoulos, N. *et al.* (2000). "Ambient Air Quality, Pollutant Dispersion, and Transport Models," *International Journal of Environment and Pollution*, European Topic Center on Air Quality Report #19/1996.

Neidell, M. J. (2004). "Air Pollution, Health, and Socio-Economic Status: The Effect of Outdoor Air Quality on Childhood Asthma," *Journal of Health Economics*, 23, 1209 – 1236.

Oates, W. E. (1972). *Fiscal Federalism*. New York: Harcourt, Brace Jovanovich.

Oates, W. E. and R. M. Schwab (1988). "Economic Competition among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics*, 35, 333 – 354.

Ogawa, H. and D. E. Wildasin (2009). "Think Locally, Act Locally: Spillovers, Spillbacks, and Efficient Decentralized Policymaking," *American Economic Review*, 99, 1206 – 1217.

Ostro, B. *et al*. (2018). "Assessing the Recent Estimates of the Global Burden of Disease for Ambient Air Pollution: Methodological Changes and Implications for Low- and Middle-Income Countries," *Environmental Research*, 166, 713 – 725.

Pun, V. C., J. Manjouride, and H. Suh. (2016). "Association of Ambient Air Pollution with Depressive and Anxiety Symptoms in Older Adults: Results from the NSHAP Study, Environmental Health Perspectives," working paper.

Rangel, M. A. and T. Vogl (2019). "Agricultural Fires and Health at Birth," *Review of Economics and Statistics*, 101, 616 – 630.

Rosenthal, S. S. and W. C. Strange (2003). "Geography, Industrial Organization, and Agglomeration," *The Review of Economics and Statistics*, 85, 377 – 393.

Rosenthal, S. S. and W. C. Strange (2008). "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64, 373 – 389.

Schlenker, W. and W. R. Walker (2016). "Airports, Air Pollution, and Contemporaneous Health," *Review of Economic Studies*, 83, 768 – 809.

Seigneur, C. and R. Dennis (2011). "Atmospheric Modeling" in *Technical Challenges of Multipollutant Air Quality Management*, London: Springer Dordecht Heidelberg.

Seigneur, C. and M. Moran (2004). "Chemical Transport Models" in *Particulate Matter Science for Policy Makers*, Cambridge: Cambridge University Press.

Sheldon, T. L. and C. Sankaran (2017). "The Impact of Indonesian Forest Fires on Singaporean Pollution and Health," *American Economic Review: Papers & Proceedings*, 107, 526 – 529.

Sigman, H. (2002). "International Spillovers and Water Quality in Rivers: Do Countries Free Ride?" *American Economic Review*, 92, 1152 – 1159.

Sigman, H. (2005). "Transboundary Spillovers and Decentralization of Environmental Policies," *Journal of Environmental Economics and Management*, 50, 82 – 101.

Stock, J. H. and M. Yogo. (2005). "Testing for Weak Instruments in Linear IV Regression," in *Andrews DWK Identification and Inference for Econometric Models*. New York: Cambridge University Press, 80 – 108.

Syverson, C. (2011). "What Determines Productivity," *Journal of Economic Literature*, 49, 326 – 365.

Williams, R. C. (2012). "Growing State–Federal Conflicts in Environmental Policy: The Role of Market-Based Regulation," *Journal of Public Economics*, 96, 1092 – 1099.

World Health Organization. (2006). *WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide: Global Update 2005*, Geneva, Switzerland.

Xu, C. (2011). "The Fundamental Institutions of China's Reform and Development," *Journal of Economic Literature*, 49, 1076 – 1151.

Zhang, P., J. Zhang, O. Deschênes, and K. Meng (2018). "Temperature Effects on Productivity and Factor Reallocation: Evidence from a Half Million Chinese Manufacturing Plants," *Journal of Environmental Economics and Management*, 88, 1 – 17.

Zheng, S., J. Cao, M. Kahn, and C. Sun (2014). "Real Estate Valuation and Cross-Boundary Air Pollution Externalities: Evidence from Chinese Cities," *Journal of Real Estate Finance Economics*, 48, 398 – 414.

Figure 1: Example of wind directions between nearby and focal city included in pollution decay function estimation

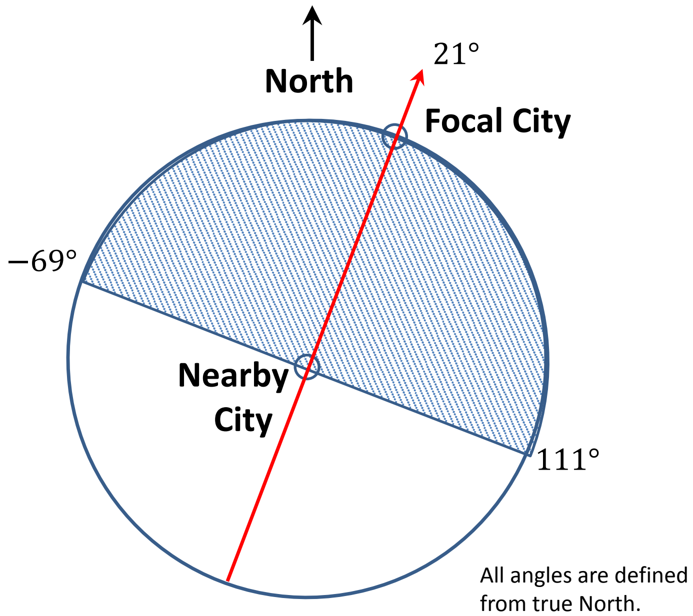


Figure 2: Example of wind directions included in estimating the causal effects of pollution on productivity (middle funnel)

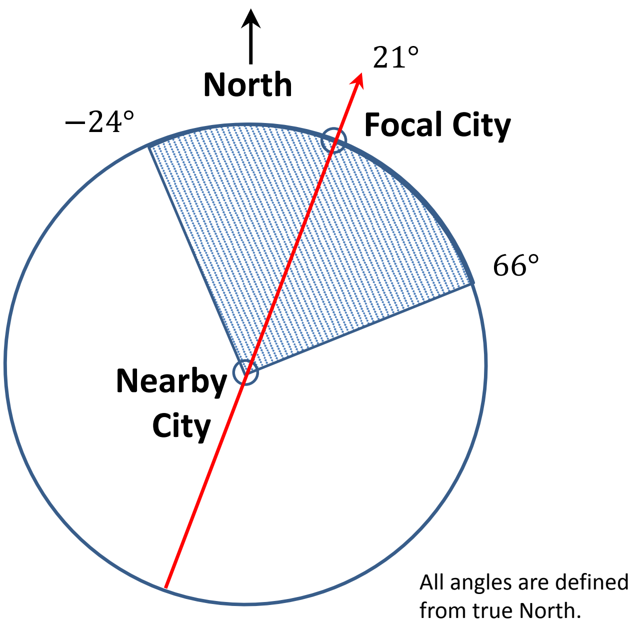
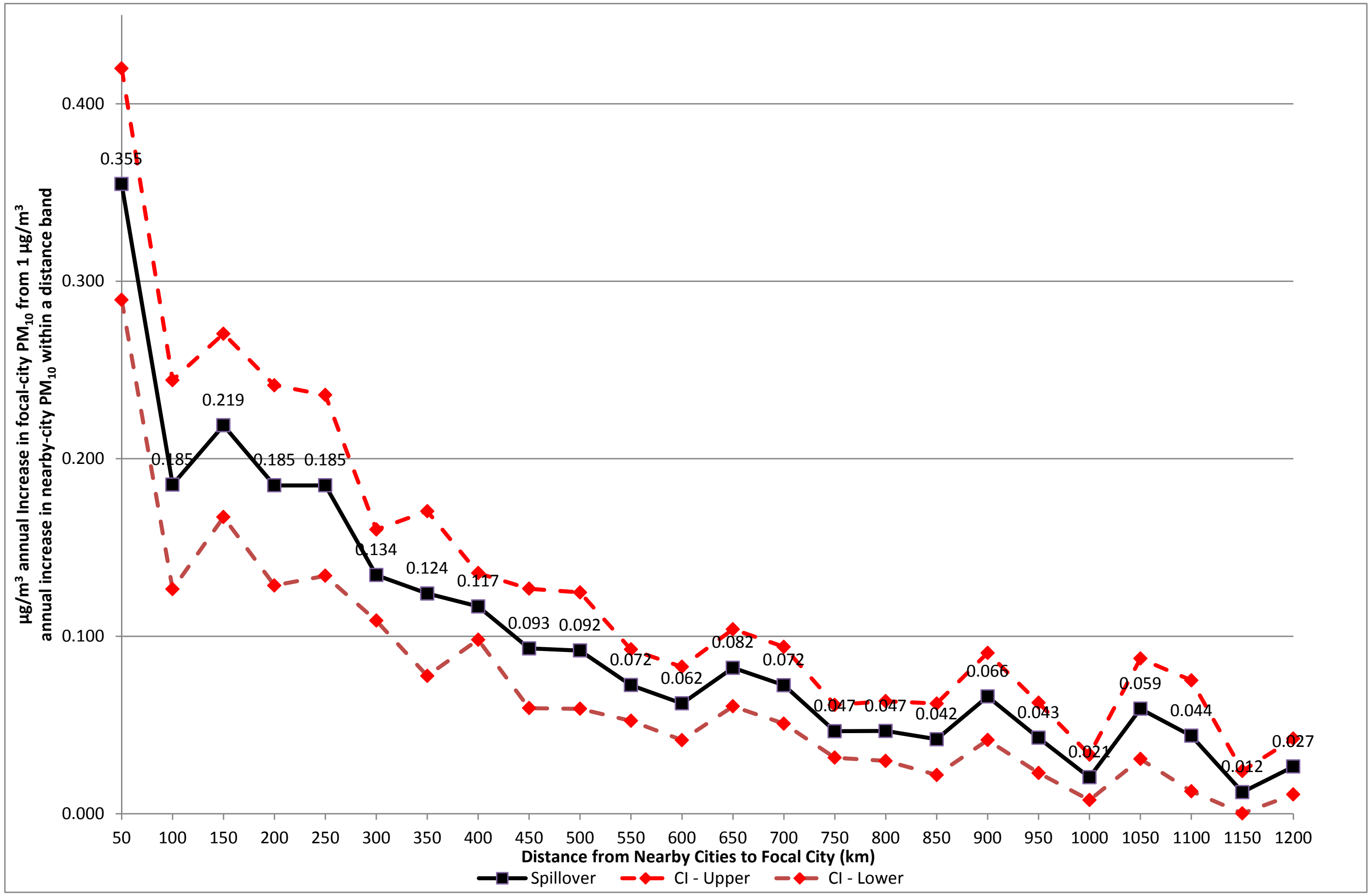
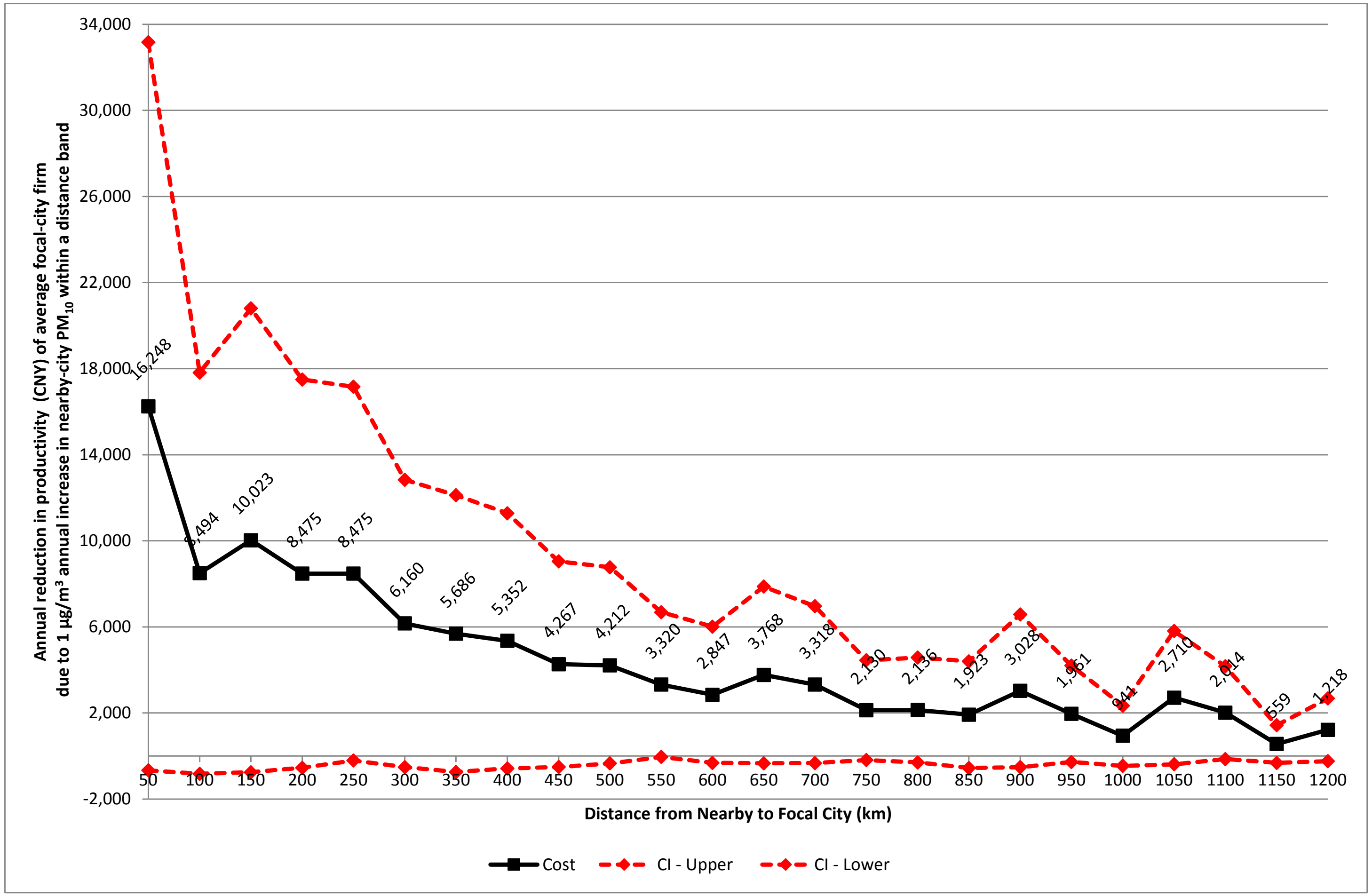


Figure 3: Pollution decay function: effect of one μg/m³ annual increase in nearby-city PM_{10} within a distance band on annual focal-city PM_{10} as a function of distance



Solid, black line shows effect of a one μg/m³ annual increase in nearby-city PM_{10} within a distance band (holding pollution in all other distance bands constant) on annual focal-city PM_{10} as a function of distance controlling for weather variables, focal-city fixed effects, and region-by-year-by-month fixed effects. Estimation allows for piecewise linear effects in increments of 50 kilometers. Effects are adjusted for the empirical distribution of wind directions during the year. Dashed, red lines show 95% confidence intervals estimated using 100 iterations of a block bootstrap by focal city.

Figure 4: Air pollution spillover effects from a one μg/m³ annual increase in nearby-city PM_{10} within a distance band on average annual labor productivity of focal-city firms as a function of distance



Solid, black line shows effect of a one μg/m³ annual increase in nearby-city PM_{10} within a distance band (holding pollution in all other distance bands constant) on average annual productivity of focal-city firms as a function of distance estimated by the two-step procedure described in the text. Estimation allows for piecewise linear effects in increments of 50 kilometers. Effects are adjusted for the empirical distribution of wind directions during the year. Dashed, red lines show 95% confidence intervals estimated using 100 iterations of a block bootstrap by focal city.

Table 1: Summary statistics for pollution decay function estimation 2001 to 2007 (N = 988,320)

	(1) Mean	(2) Std. dev.	(3) Min	(4) Max
Focal city PM_{10} ($\mu g/m^3$)	97.5	59.5	8.0	600.0
Distance between focal/nearby city (km)	1,003.9	444.0	44.0	1,799.2
Nearby cities per focal city	43.1	11.8	2.0	56.0
Fraction of days wind toward focal city		52.1%		
Fraction of days API = PM_{10}		91.9%		
# of focal/nearby cities		60		
# of focal-nearby city-year pairs		2,586		

Table 2: Summary statistics for M2SLS estimation 2001 to 2007 (150- and 300-kilometer maximum distances)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max
	150 kilometers proximity				**300 kilometers proximity**			
First-stage sample (firm-day)								
	(N = 16,271,706)				**(N = 19,339,917)**			
Focal city PM_{10} ($\mu g/m^3$)	111.6	69.0	10.0	600.0	110.5	67.8	10.0	600.0
Nearby city PM_{10} ($\mu g/m^3$)	97.5	65.2	11.0	600.0	97.2	63.2	11.0	600.0
Distance between focal/nearby city (km)	89.2	28.5	44.0	143.8	106.5	50.8	44.0	291.8
# of city-years	103				166			
# of focal cities	30				47			
Second-stage sample (firm-year)								
	(N = 243,368)				**(N = 291,339)**			
Value added (CNY1,000)	15,181.5	27,121.6	105.7	357,934.3	15,269.8	27,296.6	101.3	366,425.6
Total workers	166.9	244.7	10.0	3,012.0	171.6	252.9	10.0	3,012.0
Value added per worker (CNY1,000)	119.7	216.2	0.5	16,247.6	118.9	219.9	0.1	16,247.6
# of firms	75,390				88,716			

Summary statistics for data used in M2SLS estimation of causal effect of local air pollution on local firms' labor productivity. First-stage data is conditional on wind blowing toward the focal city.

Table 3: Causal effect of local PM_{10} on local labor productivity – OLS and M2SLS estimates using nearest-nearby city pollution within middle funnel and different maximum distances as an instrument

	(1)	(2)	(3)	(4)
	Maximum distance cutoff			
	150 km	**200 km**	**250 km**	**300 km**
Panel A: OLS (firm-year sample)				
Dependent variable:	**ln(value added/worker)**			
Mean annual focal city PM_{10}	-0.0015	-0.0003	-0.0005	-0.0005
	(0.0014)	(0.0014)	(0.0013)	(0.0013)
R^2	0.0738	0.0777	0.0740	0.0839
Sample size	243,368	264,746	276,528	291,339
Panel B: M2SLS first stage (firm-day sample)				
Dependent variable:	**Daily focal city PM_{10}**			
Daily nearby city PM_{10}	0.7172***	0.7025***	0.7004***	0.6959***
	(0.0756)	(0.0708)	(0.0687)	(0.0669)
Fraction of days wind toward focal city	0.246	0.248	0.250	0.246
KP *F*-statistic	90.0	98.4	104.0	108.1
# cities	30	40	44	47
Sample size	16,271,706	17,858,505	18,758,702	19,339,917
Panel C: M2SLS second stage (firm-year sample)				
Dependent variable:	**ln(value added/worker)**			
Mean annual predicted focal city PM_{10}	-0.0019	-0.0026*	-0.0034**	-0.0030**
	(0.0015)	(0.0014)	(0.0015)	(0.0014)
Implied elasticity	-0.198	-0.289	-0.351	-0.312
# firms	75,390	82,714	86,941	88,716
Sample size	243,368	264,746	276,528	291,339

Data included in Panel A corresponds to firm-year data included in Panel C. First stage models include firm and region-by-year-by-month fixed effects; linear and quadratic terms of daily humidity and wind speed; and categorial variables for temperature bins as described in the text. The OLS and second-stage models include firm and region-by-year fixed effects; annual averages of linear and quadratic terms of daily humidity and windspeed; and annual counts of the daily categorial variables for temperature (i.e., number of days in each temperature bin). OLS R^2 is the "within" R^2 from the fixed effects regression. Standard errors are clustered at the focal-city level in all models and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors in Panel C are also adjusted for two-stage estimation using 100 block-bootstrap iterations.

Table 4: Wald (2SLS) estimates of causal effect of local PM_{10} on local labor productivity using pollution of nearest-nearby city within middle funnel and different maximum distances as an instrument

	(1)	(2)	(3)	(4)
	Maximum distance cutoffs			
	150 km	**200 km**	**250 km**	**300 km**
Panel A: 2SLS first stage (firm-year sample)				
Dependent variable:	**Mean annual focal city PM_{10}**			
Mean annual nearby city PM_{10} (conditional on wind blowing toward focal city)	-0.2339**	-0.2680***	-0.2435***	-0.2562***
	(0.1005)	(0.0787)	(0.0786)	(0.0637)
KP *F*-statistic	5.4	11.6	9.6	16.2
# cities	30	40	44	47
Sample size	243,368	264,746	276,528	291,339
Panel B: 2SLS second stage (firm-year sample)				
Dependent variable:	**Focal city ln(value added/worker)**			
Mean annual predicted focal city PM_{10}	0.0026	0.0052	0.0080	0.0065
	(0.0038)	(0.0032)	(0.0055)	(0.0040)
# firms	75,390	82,714	86,941	88,716
Sample size	243,368	264,746	276,528	291,339

All models include include firm and region-by-year fixed effects; annual averages of linear and quadratic terms of daily humidity and windspeed; and annual counts of the daily categorial variables for temperature (i.e., number of days in each temperature bin). Standard errors are clustered at the focal-city level in all models and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors in Panel B are also adjusted for two-stage estimation.

Table 5: M2SLS estimates of causal effect of local PM_{10} on local labor productivity at different levels of aggregation in the first stage

	(1)	(2)	(3)	(4)	(5)	(6)
	Middle funnel, 300-kilometer maximum distance cutoff					
Panel A: First stage:	**Firm-Day M2SLS**	**Firm-Week M2SLS**	**Firm-Month M2SLS**	**Firm-Quarter M2SLS**	**Firm-Semi-Annual M2SLS**	**Firm-Year 2SLS**
Dependent variable:	**Focal city PM_{10}**					
Nearby city PM_{10}	0.7572***	0.7262***	0.7255***	0.5954***	0.5160***	-0.2562***
	(0.0827)	(0.0846)	(0.1133)	(0.0722)	(0.0633)	(0.0637)
Fraction of days wind toward focal city	0.246	0.246	0.246	0.246	0.246	0.246
KP *F*-statistic	83.9	73.6	41.0	68.0	66.5	16.2
# cities	47	47	47	47	47	47
Sample size	19,339,917	9,190,704	3,182,582	1,162,124	582,678	291,339
Panel B: second stage (firm-year sample)						
Dependent variable:	**ln(value added/worker)**					
Mean annual predicted focal city PM_{10}	-0.0021*	-0.0024	-0.0017	-0.0020	-0.0024	0.0065
	(0.0012)	(0.0015)	(0.0015)	(0.0017)	(0.0019)	(0.0040)
# firms	88,716	88,716	88,716	88,716	88,716	88,716
Sample size	291,339	291,339	291,339	291,339	291,339	291,339

All columns use the middle funnel in choosing days when wind blows toward focal city and 300-kilometer radius and exclude days when API is between 95 and 105. Columns 1 through 5 use M2SLS to estimate at different levels of aggregation in the first stage: daily in Column 1, weekly in Column 2, monthly in Column 3, quarterly in Column 4, and semi-annually in Column 5 - and data at the annual level in the second stage. Column 6 estimates using Wald 2SLS with data at the annual level in both stages. First-stage models include firm and region-by-year fixed effects; linear and quadratic terms of daily humidity and wind speed; and categorial variables for temperature bins as described in the text aggregated to the corresponding level. Second-stage models include firm and region-by-year fixed effects; annual averages of linear and quadratic terms of daily humidity and windspeed; and annual counts of the daily categorial variables for temperature (i.e., number of days in each temperature bin). Standard errors are clustered at the focal-city level in all models and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Second-stage standard errors are also adjusted for two-stage estimation. In Columns 1 through 5 this is done using 100 block-bootstrap iterations.