Journal of Environmental Economics and Management 118 (2023) 102767

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

Dirty density: Air quality and the density of American cities[☆]

Felipe Carozzi [*], Sefi Roth

Department of Geography and the Environment, London School of Economics, UK

ARTICLE INFO

JEL classification:
Q53
R11
I10

Keywords:
Air pollution
Cities
Density
Health

ABSTRACT

We study the effect of urban density on the exposure of city dwellers to air pollution using data from the United States urban system. Exploiting geological features to instrument for density, we find an economically and statistically significant pollution-density elasticity of 0.14. We assess the health implications of these estimates and find that increased density in an average city leads to sizeable mortality costs. Our findings highlight the possible trade-off between reducing global greenhouse gas emissions, which is associated with denser cities according to prior empirical research, and preserving local air quality and human health within cities.

1. Introduction

As of 2018, 55% of the world's population lived in urban areas, with this figure projected to reach 68% by 2050 (United Nations, 2018). A large literature examines the consequences of urbanization and provides strong evidence that the high population densities associated with urban living lead to productivity-enhancing agglomeration effects (Melo et al., 2009; Combes and Gobillon, 2015). However, an increase in density is also associated with congestion forces such as crime, higher rental prices and possibly higher levels of air pollution (see Glaeser and Sacerdote, 1999; Combes et al., 2019 and the survey in Ahlfeldt and Pietrostefani, 2019).

In this paper, we study how urban density affects pollution exposure in residential locations. We analyze this relationship empirically by producing quantitative estimates of the elasticity of Particulate Matter (PM2.5) to population density for US cities. To do so, we construct a novel data set that combines satellite-derived measures of PM2.5 concentration with administrative data on population density for the contiguous United States. This dataset allows us to conduct our analysis at different spatial scales, yielding both between- and within-city estimates of the density pollution elasticity.

Studying the empirical relationship between urban density and ambient pollution is of particular importance for three reasons. Firstly, pollution has substantial adverse effects on human health and wellbeing. The epidemiological and economic literatures document a strong link between air pollution and various health outcomes such as life expectancy, infant mortality and emergency room visits (Dockery et al., 1993; Pope et al., 1995; Chay and Greenstone, 2003; Schlenker and Walker, 2016). World bank estimates put the global health cost of PM2.5 pollution in 2019 at USD 8.1 trillion which is equivalent to 6.1% of global GDP (World Bank, 2022). A growing body of literature shows that air pollution also affects other aspects of human life such as labor productivity,

☆ We thank Gabriel Ahlfeldt, Tatyana Deryugina, Vernon Henderson, Henry Overman, Ignacio Sarmiento Barbieri and Olmo Silva, as well as seminar participants at IEB - Universidad de Barcelona, the 2018 Annual SERC-CEP conference, the LSE Annual Workshop in Environmental Economics, the 13th Meeting of the Urban Economics Association, the 2019 AERE Summer Conference and the 34th annual Congress of the European Economic Association for helpful comments and suggestions. Excellent research assistance was provided by Floris Leijten, Jozef Masseroli, and Marguerite Obolensky. Any remaining errors are our own. The research was supported by LSE, UK Cities' seed fund.

* Correspondence to: The London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK.
E-mail addresses: F.Carozzi@lse.ac.uk (F. Carozzi), S.J.Roth@lse.ac.uk (S. Roth).

https://doi.org/10.1016/j.jeem.2022.102767
Received 13 May 2021
Available online 12 December 2022

educational outcomes and crime (Graff Zivin and Neidell, 2012; Ebenstein et al., 2016; Bondy et al., 2020) suggesting that the total cost of air pollution is likely to be even higher.

Secondly, quantifying the strength of the relationship between density and ambient pollution is policy relevant. Urban density is largely shaped by urban planning decisions: Land markets are heavily regulated by policies that are often explicitly designed to generate specific densities both in neighborhoods within cities and in cities as a whole. Examples include the use of density zoning – the specification of densities of development in a city's zoning ordinance – floor-to-area and maximum height limits on new buildings, minimum lot size restrictions, and growth boundaries to contain urban sprawl, to name a few. Because these policies are widespread both inside and outside of the United States, understanding the impact of density on air pollution helps us understand the costs and benefits of policies that promote or hinder densification.

Finally, our empirical analysis is motivated by the fact that the sign of the effect of urban density on air pollution exposure is theoretically ambiguous. Driving is an important contributor to urban air pollution and previous empirical studies show that dense locations are characterized by lower levels of vehicles miles traveled (see e.g., Duranton and Turner, 2018; Boarnet and Wang, 2019; Ihlanfeldt, 2020). Other work indicates that density can induce lower levels of emissions per capita from other sources such as heating (see e.g., Glaeser and Kahn, 2010; Borck and Brueckner, 2018). Unfortunately, the physical effect of density on concentration can offset these reductions in per capita emissions. For example, dense urban living reduces the average distance between individuals, limiting spatial decay of emissions between polluters. Thus, for a fixed level of emissions per capita, increasing density can result in higher pollution exposure. The overall effect of density on pollution will depend on whether the reduction in emissions per capita dominates the increase in pollution from the spatial concentration of emitters or vice versa, and that is an empirical question. In the process of answering this question, we produce quantitative estimates of density-pollution elasticities for the United States urban system, which themselves constitute one of our main contributions.

Estimating the effect of urban density on air pollution is empirically challenging for several reasons, including the presence of unobserved correlated factors and reverse causality. More specifically, population densities are not randomly assigned as residents sort themselves into areas based on various characteristics including local amenities and employment opportunities. For example, many productive activities (e.g., manufacturing) generate pollution; if residents sort themselves into areas close to these activities, a naive OLS estimation may overstate the true effect of densities on pollution. Furthermore, empirical evidence demonstrates that pollution itself affects location decisions directly, resulting in a reverse causality problem which would bias our coefficients towards zero (see Heblich et al., 2021).

We overcome these empirical challenges by using an instrumental variable (IV) strategy, with instruments inducing exogenous variation in density without affecting pollution directly. For this purpose, we use three variables capturing geological features of US urban areas: earthquake risks, the presence of aquifers in and around urban areas and soil drainage capacity.[1] We complement our main empirical approach with estimates obtained using a traditional long-lag instrument, which measures urban population density in the 1880 US Census (similar to Ciccone and Hall (1996) and subsequent work) and a fixed-effect specification based on a two-period panel covering the years 2000 and 2010.

We find a positive and statistically significant relationship between city-level population density and residential exposure to ambient air pollution, which we define as population-weighted average pollution concentration. Our preferred instrumental variable estimates suggest that a doubling of density – which is equivalent to changing population density in Houston to that of Chicago – increases PM2.5 exposure by 0.73 ($\mu g/m^3$). This effect, roughly equal to two-fifths of a standard deviation, is large given the substantial variation in densities between urban areas in the United States. We also estimate an exposure-density elasticity of 0.14, indicating that a 1% increase in density increases average residential PM2.5 exposure by 0.14%. To put this in perspective, according to a recent survey of the quantitative literature (Ahlfeldt and Pietrostefani, 2019), the density elasticity of wages and energy use reduction are 0.04 and 0.07, respectively. Reassuringly, our results using alternative estimation strategies based on panel variation in density, a historical population density instrument, and monitoring station data to measure pollution yield similar qualitative findings to those in our main results.

We conduct several complementary analyses to provide suggestive evidence on the mechanism behind these findings. First, we use different specifications to explore the impact of scale effects associated with a city's population on the link between pollution and density. Second, we use data on the location of polluting industries to study whether local differences in sectoral composition can explain our main results. Results from both exercises indicate limited evidence for these two channels. We also use our within-city analysis to show that denser locations within cities are indeed associated with higher pollution levels. This is consistent with the notion that the effect of city-level measures of density on average pollution exposure can result from the reduced distance between sources and affected households due to higher spatial concentration of emitters in denser urban areas.

The effect of density on local air pollution exposure is only economically relevant if it translates into significant health and well-being effects on inhabitants. We evaluate this by computing the pollution-induced mortality impacts of density following a similar analytical strategy to the one taken by the US Environmental Protection Agency (EPA) in their Regulatory Impact Analysis.[2] Using concentration–response functions from the literature and the EPA official figure for the Value of Statistical Life (VSL), combined with our main between-city IV estimates, we find that a doubling of population density is associated with an increase in annual mortality costs of as much as $630 per capita. This is a large cost that highlights the importance of incorporating air quality when

[1] Geological instruments for density were initially proposed in Rosenthal and Strange (2008) and have often been used in subsequent work in the agglomeration literature (see for example Combes et al., 2019).

[2] Regulatory Impact Analysis (RIA) is used by the EPA to support the development of national air pollution regulations as required under the 12 866 and 13 563 Executive Orders.

considering the consequences of suburbanization. Finally, as previous research suggests that compact cities are linked with lower greenhouse gas emissions, our results indicate a potential trade-off in the environmental implications of dense urban development, with increased densities associated with lower global externalities (greenhouse gas emissions) but increased local externalities (local air pollution).[3] We conduct back-of-the-envelope calculations to evaluate the net environmental impact of population density. We find that the local air pollution costs far outweigh the global environmental benefits of density in terms of reduced CO2 emissions.

Our study provides credible estimates of the causal relationship between population density and air pollution for the United States. The paper closest to our work is Borck and Schrauth (2021), which studies the effect of district-level density on urban air pollution in Germany using empirical strategies similar to those employed here. We see both studies as complementary, and we compare both sets of results in Section 4. That said, our paper differs in several ways. First, our analysis is based on data from the entire United States which provides an opportunity to estimate this relationship across a larger and more diverse urban system. Second, we provide complementary estimates from a within-city analysis that quantifies the effect of local changes in density on local pollutant concentration. Finally, we explicitly estimate the costs of densification in terms of mortality.

Several other studies estimate the relationship between air pollution and density using partial correlations in different contexts. In their review on the economics of density, Ahlfeldt and Pietrostefani (2019) use data on over 300 functional urban areas from OECD countries to estimate an elasticity of particulate matter concentration to population density of 0.12. Focusing on the US, Albouy and Stuart (2014) report an elasticity of 0.15 and Clark et al. (2011) find that a 1 interquartile range (IQR) increase in population density is associated with a 0.36-IQR increase in PM2.5 population-weighted concentration. Chen et al. (2020) studying Chinese cities report a cross-sectional elasticity of 0.41, though they report a negative elasticity of −0.26 when controlling for city fixed effects in a panel specification. Our paper improves over this section of the literature by introducing an empirical strategy that is better suited to deal with issues of causality in this context. For example, our instrumentation strategy avoids issues of reverse causality arising from the effect of pollution on household location decisions.

Our study also provides a valuable insight for the design of public policy and urban planning in particular. Our results emphasize that, despite the usual claim that denser cities tend to be more environmentally friendly (in terms of lower greenhouse gas emissions), residential exposure to air pollution is higher in denser cities. This means that there may be a trade-off between potential reductions in emissions and local increases in pollutant concentration associated with changes in density. Furthermore, our elasticities and cost estimates can aid city planners in performing cost benefit analysis of different densification policies.

The rest of the paper is laid out as follows. In Section 2 we describe our data. In Section 3 we explain our empirical strategy. In Section 4 we present our main results and robustness tests. In Section 5 we discuss these results and possible mechanisms explaining our findings. In Section 6 we calculate the mortality impacts and economic costs of air pollution-induced by density, based on our estimates. We conclude in Section 7.

2. Data and descriptives

For our empirical analysis, we combine information on air pollution concentration, population density, demographics and geological features from several sources. We begin by constructing our dataset as a 0.01×0.01 degrees grid over the conterminous United States territory. This grid is based on a raster of average PM2.5 concentration measurements obtained by combining the Aerosol Optical Depth retrieval from the NASA MODIS instrument adjusted using ground-level monitoring station level data as detailed in Van Donkelaar et al. (2015). The average cell size is approximately 1 square kilometer and the inter-quartile range in cell size is only 0.1 sq. km. Throughout most of our analysis, our dependent variable measures PM2.5 average yearly concentration for 2010.

We focus on PM2.5, an heterogeneous mixture of tiny particles (less than 2.5 μm in diameter) from various sources, as it poses the greatest risk to human health (Tschofen et al., 2019; Environmental Protection Agency, 2022). According to the World Health organization and the World Bank, most air pollution-related deaths are linked to PM2.5 (see World Health Organization, 2022; World Bank, 2022). As such, reducing PM2.5 has also become the main objective for many legislative initiatives around the world including the Zero Pollution Action Plan in the EU and the UK Environment Act.

We use population data from the US census for 2010 by spatially matching our grid cells with census-blocks.[4] We then employ this information to compute the grid cell level and city level density measures. We additionally incorporate into our dataset information on demographic characteristics of local populations including income and education proxies, population age, and housing tenure.

To obtain our main instrumental variable estimates, we use three different instruments that may affect population density but not air pollution directly. We obtain variables measuring earthquake risks and the presence of aquifers from the United States Geological

[3] A well-developed empirical literature studies the effect of density on emissions (rather than pollution concentration). Zheng et al. (2011) show that emissions per capita are negatively associated with density using data from 74 major cities in China. Cirilli and Veneri (2014) found that more compact areas in Italy are linked with lower levels of CO2 emissions per commuter. Similarly, de Thé et al. (2021) present evidence that more compact cities have lower car emissions per household. de Thé et al. (2021) and Carantino et al. (2018) find a negative effect of density on household fuel consumption across US cities and in France, respectively. Finally, Sarzynski (2012) and Castells-Quintana et al. (2021) show that denser cities are correlated with lower emissions of air pollutants per capita using a cross-section of world cities. Theoretical treatments of the relationship between emissions and urban form in a system of cities can be found in Gaigné et al. (2012), Denant-Boemont et al. (2018), Borck and Tabuchi (2019) and Borck and Pflüger (2019). Borck and Tabuchi (2019) discuss the trade-off between local and global pollution when studying optimal city sizes in a city system model.

[4] Imputation of population data to our grid cells is performed by assuming a uniform spatial distribution of populations within the census block, overlaying our polygon grid on census block-group shapefiles distributed by the US Census Bureau, and aggregating data back to our grid cells using spatial weights computed using surface areas.

Table 1
Descriptive statistics.

	Mean	Std. dev.
A. Spatial cells		
PM2.5 average (satellite data)	5.90	2.54
Population Density	63.35	334.48
Earthquake risk (3 cat.)	0.69	0.47
Aquifers (2 cat.)	0.28	0.45
Population density in 1880	8.95	66.55
Minimum dist. to water (km)	55.62	49.42
Latitude	38.5	5.11
Longitude	−98.7	14.82
Gridcell area	1.0	0.07
Observations	4,356,408	
	Mean	Std. dev.
B. Cities		
PM2.5 spatial average (satellite)	6.98	2.21
PM2.5 residential-weighted (satellite)	7.84	1.94
PM2.5 (monitoring stations)[a]	9.11	2.75
Population Density	55.08	78.47
Earthquake risk (3 cat.)	0.61	0.48
Aquifers (2 cat.)	0.28	0.39
Population density in 1880	11.54	12.96
Minimum dist. to water (km)	53.40	42.43
Latitude	38.0	4.92
Longitude	−91.8	13.00
Gridcell area	1.0	0.07
Observations	933	

Notes: Descriptive statistics for our within and between city samples. Panel A presents mean and standard deviation for a set of key variables of interest. Panel B presents statistics for these variables after aggregating at the city (CBSA) level.
[a]Only 546 cities have $PM_{2.5}$ monitoring station data.

Survey (USGS) (also used in Duranton and Turner, 2018), and data on soil drainage quality from NRCS State Soil Geographic Data Base. We match our grid cells to the geological data using grid cell centroids to spatially impute data on aquifers, earthquake risks and soil drainage quality. An alternative instrument, used in our robustness checks, is population density data at the county level obtained from the 1880 United States census. We impute this data on our grid cells using spatial matching based on the assumption of uniform population distribution within 1880 counties.[5]

Our main analysis is based on between city comparisons, where the definition of city is based on the Core-Based Statistical Areas (CBSAs) as defined by the US Office of Management and Budget. We include both Metropolitan and Micropolitan statistical areas in most of our city-level analysis. These areas are defined as an aggregation of counties based on commuting patterns around an urban core. In using these areas, we attempt to approximate a working definition of a functional urban area. The definition of CBSAs used in this paper is based on the 2010 Census, with the associated shapefiles obtained from the US Census Bureau. We complement our dataset with gridded data on PM2.5 concentrations from ground-level monitoring stations (obtained from EPA AirData) and industrial composition at the county level from the County Business Patterns (CBP) dataset. We use the ground-level monitoring station data to validate our satellite-derived pollution measures. The scatter plot in Appendix Figure A1 shows the correlation between monitoring station measures of PM2.5 concentration and concentration measures obtained using our satellite-derived data. In both cases local measures are aggregated to the city level. The correlation is high, as expected, standing at roughly 80%.[6]

The dataset assembly process is illustrated using the metropolitan area (MSA) of Minneapolis-St. Paul-Bloomington in Appendix Figure A2. Top panel A shows a satellite photograph of the Twin Cities, with the Mississippi crossing the urbanized area from north-west to south-east. Panel B presents our pollution raster, with darker shades indicating higher PM2.5 concentration levels. Points in panel B indicate the location of AirData measuring stations in the area, which we will use in our robustness checks. Finally, bottom panel C indicates the resolution of our population data, with polygons indicating census block groups. We spatially impute data to our grid cells and then aggregate to CBSAs to obtain our city-level dataset. Table 1 provides descriptive statistics for the main variables of interest across our 4,356,408 spatial cells (Panel A) and our aggregated dataset of 933 cities (Panel B). Panel B shows that there is substantial variation in both density and air pollution levels across the US urban system. We can also see that population densities of US cities have increased almost fivefold since the 19th century. Importantly, annual mean PM2.5 is lower

[5] Note that, while the assumption of uniform distribution is clearly a simplification that could lead to measurement error, this should not have a substantial impact on our main estimates. Measurement error in the instruments could affect their relevance but should not generate bias in the coefficients of interest unless the measurement error itself is correlated with pollution concentration.

[6] Sullivan and Krupnick (2018) use the same satellite-derived source of particulate matter concentration we use in this paper to complete the large gaps in the coverage of the pollution monitor network in the United States.

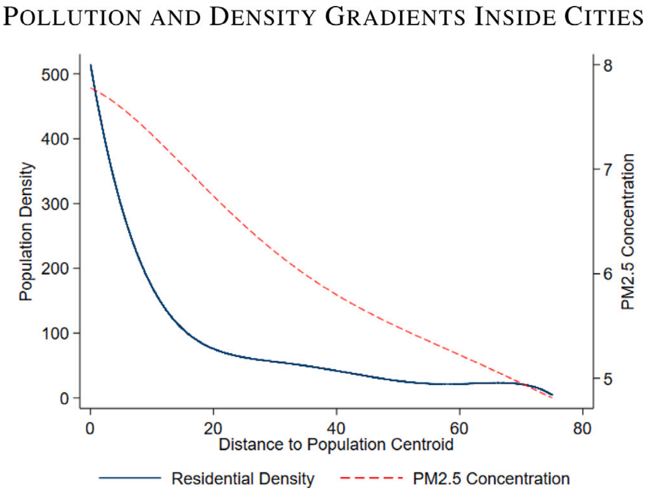

Fig. 1. Pollution and Density Gradients Inside Cities. *Notes*: Horizontal axis represents to distance to the CBSA population-weighted centroid. Vertical axes correspond to population density (left axis) and satellite-derived PM2.5 concentration (right axis). Lines obtained by estimating 5th degree polynomials over grid-level data. Blue line corresponds to population density and red line to PM2.5 concentration.

for the satellite measure suggesting that ground measuring stations are located in more polluted areas. This endogenous location of measuring stations generates a measurement error problem that exists, to varying degrees, in much of the empirical literature on ambient air pollution. Moreover, PM2.5 measuring stations were altogether absent in 41% of CBSAs in 2010. Because these cities have lower average pollution levels than those with monitors, the use of data from ground monitors generates a sample selection problem that can affect our estimates. To further explore the distribution of population and pollution, we provide distance gradients for both population density and PM2.5 concentration in Fig. 1 using data for all cities. The horizontal axis represents the distance to the CBSA population centroid, and the solid line represents population density.[7] We observe the usual downward sloping density gradient that characterizes most cities (see for example McDonald, 1989, or Bertaud and Malpezzi, 2003). Interestingly, we observe a similar downward-sloping pattern in the dashed line representing particulate-matter pollution concentration. The fact that this line has a gentler slope is likely to be the result of the slow spatial decay of PM2.5 concentrations. Similar patterns can be observed in Appendix Figure A3, which is obtained using ground-based monitor measures of pollutants.[8]

3. Empirical strategy

3.1. Between-city analysis

The goal of our empirical analysis is to estimate the effect of urban density on residential pollution exposure and our main analysis exploits between-city variation for this purpose. The resulting estimates are relevant for city-wide planning policies and have been the focus of the limited literature on this topic. Our dependent variable in this exercise is average population-weighted PM2.5 concentration at the city (CBSA) level. This is built by aggregating data from our spatial cells into city-level observations, combining satellite measures of pollution concentration with residential population counts from the census. Formally, city c average population-weighted particulate concentration is given by the following weighted average[9]:

$$PM2.5_c^{exp} = \sum_{i=1}^{N_c} PM2.5_{c,i} \times \frac{Pop_{c,i}}{\sum_{j=1}^{N_c} Pop_{c,j}}$$

[7] The CBSA population centroid is computed by calculating the weighted average of latitude and longitude for grid-cells within a city, where weights are given by the fraction of total CBSA population in a grid-cell.

[8] Note that average PM2.5 concentrations are higher when using ground-based monitoring stations because monitoring stations tend to be located in more polluted areas. However, the use of satellite images to measure pollution might be affected by cloud coverage, which in certain circumstances can lead to measurement error. We believe that this is not a threat to identification in our context for two main reasons. First, cloud cover may affect readings at the daily or weekly levels but should have a more modest impact when considering annual averages. The high correlation between monitoring station measures of PM2.5 concentration and concentration measures obtained using our satellite-derived data is reassuring in this regard. Second, our pollution measure is our dependent variable rather than the independent variable, which means that even if we do measure pollution with error in this context, our estimates remain consistent (though less precise).

[9] The pollution exposure of individuals depends on a variety of factors including their geographical location across time and their time spent indoors/outdoors. Since we only observe the residential location, we use the above definition to measure exposure.

where i indexes grid cells within city c, N_c is the number of cells within c, $Pop_{c,i}$ is the population on grid cell i and $PM2.5_{c,i}$ is average yearly PM2.5 concentration on grid cell i.[10] Our main estimating equation regresses the natural logarithm of average residential PM2.5 exposure ($Ln(PM2.5_c^{exp})$) on the logarithm of urban population density ($Ln(Density_c)$), computed by aggregating populations and areas to the CBSA level:

$$Ln(PM2.5_c^{exp}) = \beta Ln(Density_c) + \gamma' X_c + \alpha_s + \nu_c \tag{1}$$

where X_c corresponds to our vector of controls, α_s is a state fixed-effect and β is our coefficient of interest which indicates how PM2.5 exposure increases as a result of an increase in urban density. This estimate is interpreted here as an elasticity, but we also provide estimates with the dependent variable in levels, since these are necessary to obtain a monetary measure of health costs as calculated by the EPA. We will estimate β both via OLS and using two alternative instrumental variable strategies described below.

We present results with and without state fixed effects and with two different sets of controls. Our basic set of controls includes average maximum and minimum temperatures and decimal degree coordinates of the CBSA centroid (latitude and longitude).[11] In some specifications we include an additional set of control comprising distance to water bodies, average precipitation, average soil slope, a dummy taking value 1 for coastal cities (population-weighted centroid less than 50 km of a major coastline), distance to a major coastline (ocean or great lakes) and the number of power stations in the city. The role of these controls is twofold. In the first place, some controls simultaneously affect density and pollution levels, so their omission would lead to omitted variable bias in the OLS estimates. For example, temperature may affect pollution through variation in the usage of heating systems and/or various meteorological conditions (e.g., thermally induced convection) and may affect density via its effect on population location decisions (see Glaeser et al., 2001). Secondly, controls can also be helpful when using instrumental variables (IV) to obtain exogenous variation in population density. The fact that the controls explain part of the variation in density, means that they improve the precision of first-stage estimates and, through that, the precision of 2SLS estimates of the effect of interest. We will discuss the IV strategy in detail below.

Finally, we can use our detailed spatial data on pollutant concentration to study the fraction of a city's population that resides in locations with yearly PM2.5 levels above the EPA primary national ambient air quality standard (NAAQS) for particulate pollution of 12 micrograms per cubic meter. We can directly create this variable using our gridded data on pollution concentration and residential locations. We then use this variable as an outcome in Eq. (1) to test whether the average effects from our original specification translate into more people living in areas that do not meet the NAAQS.

3.2. Within-city analysis

We complement our main between-city analysis by looking into urban areas and estimating how local changes in density affect pollutant concentration. The purpose of the analysis is to obtain a quantitative estimate of the causal effect of local density on local pollutant concentration. When looking within cities, the baseline estimating equation takes the form:

$$Ln(PM2.5_i) = \omega Ln(Density_i) + \gamma' X_i + \alpha_c + \varepsilon_i \tag{2}$$

where i indexes grid cells, $Ln(PM2.5_i)$ is the natural logarithm of satellite-derived PM2.5 concentration, $Ln(Density_i)$ is the log of grid cell level population density, α_c is a CBSA level fixed effect, and X_i is a vector of controls.[12] The CBSA level effects ensure we only exploit within-city variation. We use the same sets of controls described in the between-city analysis above. We cluster the standard errors at the city-level to account for correlated shocks within each CBSA. In this exercise, our parameter of interest is ω, which can be interpreted as an elasticity. We will also report estimates when PM2.5 concentration is kept in levels.

3.3. Instrumental variables

The causal interpretation of the coefficient of interest in both the within and between city estimates requires variation in density to be exogenous to other determinants of air pollution. Urban density is shaped by a host of factors ranging from sectoral specialization, locational amenities, access to employment opportunities and, potentially, air quality itself. This is problematic because some of these factors could very well affect pollution directly, therefore becoming confounders in the regression equations above. While controlling for other determinants of pollution or state effects may help, there is no guarantee that all confounders have been accounted for.

We overcome this problem by using an instrumental variable strategy that employs three different instruments to induce credibly exogenous variation in density. Before discussing our instruments in detail, it is useful to go through the randomization thought experiment, focusing specifically on the between city case.

---

10 Interestingly, our results are only partially affected by taking this weighted average. Using a simple mean of PM2.5 concentration within our cities has a modest impact on estimated coefficients.

11 We do not control for total urban population in our main specification but conduct a separate analysis to account for the role of city size on the link between pollution concentration and density in Section 5. Controlling for population in our main specification barely affects our coefficients of interest.

12 The gridded data from Van Donkelaar et al. (2015) does not provide precise local pollution concentrations in all grid cells. The reason is that the method developed by this group incorporates information measured at different levels of aggregation, with some information recorded at a coarser resolution. We see this as inducing a source of measurement error in the outcome which reduces precision but should not induce a bias in estimation as long as measurement error is orthogonal to density.

Ideally, one would want to randomize urban density across cities, for example by randomizing maximum height restrictions or zoning regulations. Even if this were possible – it is not – the result is likely to affect a multitude of different urban outcomes, through static effects on the city and migration within the urban system. These factors will, in turn, collectively influence urban air pollution. We argue that despite the reduced-form nature of this type of exercise (where a multitude of potential mechanisms could be in operation simultaneously), this is the policy parameter of interest. Densification policies by city governments are also likely to have a large set of impacts on urban structure and across the city system. The variation in density we use for estimation is not induced by planning policy decisions but rather by other density shifters relating to the physical and historical environment of cities. However, we expect estimates obtained from these induced variations to remain relevant for planning policy.

Our main empirical analysis uses three geological features as instruments for population density.[13] More specifically, we use the fraction of the urban footprint with aquifer presence, a measure of average earthquake risks and an estimate of soil drainage quality. The rationale for the aquifer variable is that new dwellings in the periphery of urban areas need either to connect with the municipal network or to directly connect with an underwater source. As the cost of connecting to the municipal network is increasing in the distance to other connected dwellings and the option of the underwater source is only available if there is an aquifer where the dwelling is located, cities with more land over aquifers can sprawl out further and contain more sparse development at lower densities. This instrument is motivated by the work in Burchfield et al. (2006) which reports that aquifers in the urban fringe are associated with urban sprawl.

The rationale for our earthquake risk instrument is the expectation that the risk of an earthquake might influence building regulations, construction practices and the space between buildings, thus also affecting urban density.

Finally, the soil drainage quality variable is expected to affect land suitability for building at different densities. In fully urbanized land, a significant fraction of rainfall is drained through drainage networks and sewage systems (Konrad, 2003). However, at lower densities, soil drainage capacity is important to avoid stagnant water and, possibly, floods. In addition, because high drainage soils are composed of relatively large particles, which leave substantial empty spaces between them, it is not ideal for the laying down of heavy infrastructure, thus penalizing high-density development.

The use of these geological instruments requires making two assumptions. The first is that the instruments do not affect the outcome directly (exclusion restriction). We think this assumption is likely to be satisfied in practice for all three instruments. Neither the presence of an aquifer nor earthquake risk are bound to themselves affect the outcome except through their influence on human settlement. Aquifers are underground and earthquakes are naturally rare events and therefore earthquake risk is unlikely to affect average yearly pollutant concentration. While soil features can occasionally induce particulate matter pollution in the form of dust, this is usually limited to relatively large particles (e.g., PM10) rather than the smaller particles that characterize PM2.5.

A second assumption that is necessary for instrument validity is that instruments are not correlated with omitted variables that could directly affect PM2.5 and density (orthogonality). In principle, our instruments could be correlated with potential confounders. For example, earthquake risk is disproportionately concentrated in relatively rugged locations of the United States, where terrain makes urban development more difficult. Nevertheless, we believe that this is very unlikely that all our instruments violate the orthogonality assumption, especially as our different sets of controls and state effects should help in dealing with these confounders. Reassuringly, the estimates obtained for different specifications are very similar, suggesting that the orthogonality assumption is satisfied in our case.

As a complement to the empirical strategy above, we introduce an alternative instrumental variable based on historical population as recorded in the 1880 US census. This period took place before much of the technological revolutions in transportation that affect contemporaneous air pollution and would also precede more recent patterns of industrial location. Moreover, because of the persistence of buildings and other urban features, late XIX century density could affect urban density today. Therefore, we expect the relevance condition to be satisfied. The use of historical population instruments for density was initially proposed by Ciccone and Hall (1996) and has often been used in the literature on agglomeration economies since (see for example Combes et al., 2008).[14]

Formally, our IV estimates for the between-city analysis are obtained following the standard two-stage least squares procedure (2SLS) as follows:

$$Ln(Density_c) = \delta' Z_c + \gamma_z' X_c + \theta_s + u_c \tag{3}$$

$$Ln(PM2.5_c^{exp}) = \beta Ln(\widehat{Density_c}) + \gamma' X_c + \alpha_s + v_c \tag{4}$$

where Z_c is our vector of instruments and δ represents a conformable vector of first-stage coefficients. All other variables defined as above. Throughout our analysis, we provide results for different specifications, including or excluding controls and state fixed effects. We follow the same strategy to obtain our within-city estimates, replacing grid cells as the unit of observation and controlling for CBSA effects.

[13] A similar strategy is followed in Rosenthal and Strange (2008), Combes et al. (2010) and Duranton and Turner (2018).

[14] A description of the intuition behind both the population lag and geological instruments and their limitations can be found in Combes and Gobillon (2015).

PM2.5 Exposure v. Density Scatter

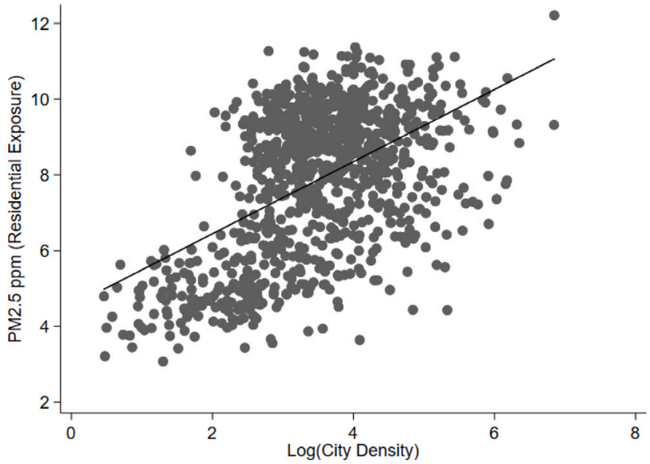

Fig. 2. PM2.5 Exposure v. Density Scatter. *Notes:* Vertical axis represents PM2.5 average residential exposure (in μg /m^3), as obtained from the satellite-derived measures. Horizontal axis represents the natural logarithm of population density. The points represent 933 CBSAs (metro and micropolitan areas). The black line is estimated by OLS using the underlying data.

4. Results

4.1. Between-city estimates

As a preview to our baseline results, Fig. 2 provides graphical evidence on the cross-sectional relationship between city-level particulate concentration and population density where the vertical axis measures $PM2.5_c^{exp}$ and the horizontal axis measures the logarithm of density at the city level. Clearly, there is a strong positive relationship between both, as indicated by the regression line overlaid on the figure.[15] One potential concern when interpreting this figure is that denser cities may have higher levels of air pollution exposure, not because they are dense, but rather because they are populous. To illustrate that the urban scale is not driving the correlation observed in Fig. 2, we regress $PM2.5_c^{exp}$ on a fourth-degree polynomial of population. We then obtain the residuals of this regression and plot them against log density. The corresponding scatter plot is provided in Appendix Figure A5. We observe the relationship between $PM2.5_c^{exp}$ and density is largely preserved after this procedure, indicating that the observed correlation is not driven mechanically by city size. We conduct a more detailed analysis of the role of city-scale in Section 5.

Table 2 provides baseline estimates of Eq. (2), obtained using ordinary least squares. The top panel of the table provides estimates of the elasticity of pollution with respect to population density. The bottom panel reports coefficients of a specification in which the PM2.5 concentration variable is kept in levels. Different specifications, displayed in columns, adding fixed effects and additional controls as indicated in the table. In all cases, our unit of observation is the CBSA. Our preferred estimates are those for which the full set of controls and state fixed effects are included in the estimating equation. Focusing on the top panel, our baseline results indicate an elasticity of 0.073, significant at the one percent level. This suggests that a 1% increase in density would result in a 0.07% increase in average residential PM2.5 exposure. As discussed above, these estimates can be biased by confounders or reverse causality. Specifically, air pollution may lead to lower equilibrium densities if households sort spatially in response to it. Recent evidence of air pollution affecting household location decisions can be found in Heblich et al. (2021), Khanna et al. (2021) and Chen et al. (2022).

Before proceeding to the IV results, we first provide estimates of our first-stage regression in Appendix Table A1. These result from estimating Eq. (3), using our geological variables as instruments. We observe that across specifications our instruments are jointly significant, as indicated by the F-statistics reported in the table foot, which lies consistently around 20. Both the aquifer and soil drainage instruments have the expected signs, given that both aquifer presence and high-quality soil drainage predict low-density development. Note that the expected sign for the aquifer instrument is different for the between and within city analysis. In the between-city analysis, we expect the sign of the first stage coefficients to be negative, as cities with more land over aquifers can sprawl out further and contain more sparse development and lower densities. In contrast, we expect to see a positive sign in the within-city analysis as residents are likely to sort into locations where aquifers are present. We will show that this is indeed the case when we present our within-city analysis below.

[15] A non-parametric depiction of this relationship using population density bins is provided in Appendix Figure A4 in the Appendix.

Table 2
Between-city baseline OLS estimates.

	Log(PM2.5) - elasticities		
Log(Pop. Dens.)	0.085***	0.078***	0.070***
	(0.007)	(0.007)	(0.007)
Observations	933	933	933
	PM2.5		
Log(Pop. Dens.)	0.578***	0.535***	0.471***
	(0.054)	(0.046)	(0.047)
Full controls	No	No	Yes
State-FE	No	Yes	Yes
Obs.	933	933	933

Notes: Baseline OLS estimates. City-level regressions. For the first row of estimates, the dependent variable is the natural logarithm of PM2.5 residential exposure as defined in the text. The dependent variable for the second row of estimates is this variable in levels. All specifications include latitude, longitude and average maximum and minimum temperatures as controls. The specifications in columns 2 and 3 add state effects. The specifications in column 3 add a set of additional controls as detailed in the text. Robust standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Finally, the coefficient on earthquake risk is harder to interpret. The effect of earthquakes on the risk of collapse may not be increasing in building heights, as different buildings will have different resonances and therefore be affected differently by different types of earthquakes.

As an additional check on the suitability of our instruments, we modify equation 2 and estimate the effect of our instruments on variables measuring the presence of fossil fuel power stations, and on the Wharton index of land use regulation (Gyourko et al., 2008). In the case of power plants, the concern is that geological characteristics (e.g., earthquake risks) could affect the location of power plants across cities and these could, in turn, affect PM2.5 pollution directly. Estimates for these balancing tests are provided in Appendix Table A2. In columns 1 and 2 we use, respectively, the number of oil and power plants in each city as outcome variables. Columns 3 and 4 use dummies indicating the presence of at least one plant of each type instead. Reassuringly, the coefficients on our estimates are always insignificant at the 5% level, with only one coefficient being weakly significant in one specification. The joint significance test statistics for the instruments never reject the null of joint insignificance, with F-statistics between 0 and 2 in all columns. When considering land use regulation, the concern is that it could simultaneously be affected by air pollution and correlated with our instruments. The results for these balancing tests are reported in Appendix Table A3.[16] We can see that once we include state effects or the full set of controls (columns 2 to 4), we cannot reject the null of all instrument coefficients being jointly not significant (see F-statistic provided in the table foot). Only the earthquake risk variable remains marginally significant and this is not wholly surprising, as we would expect certain building regulations to be more stringent in areas with high earthquake risk. Having provided further evidence on the validity of our instruments in the context, we now turn to our main results.

Table 3 presents the results of our main IV estimation. The first-row reports estimated elasticities, and the second reports estimates obtained with PM2.5 exposure measured in levels. We provide results using different specifications, with our preferred specification including both state dummies and a full set of controls (third column). Our estimated elasticity is now 14% and is statistically significant at all conventional levels. This indicates that a 1% increase in density will increase average residential PM2.5 exposure by 0.14%. This is slightly larger than our baseline estimate, indicating a downward bias under ordinary least squares. This is consistent with reverse causality running from pollution concentration to population density, as discussed above.

It is worth noting that our estimated elasticity is close to the elasticity of 12.4% reported in Ahlfeldt and Pietrostefani (2019) using a different sample of cities. In Borck and Schrauth (2021) the headline estimated elasticity for PM2.5 in German districts is in the 2%–8% range, although statistically insignificant. The authors argue that the imprecise estimates are due to the recent and incomplete nature of the German network of measuring stations. Complementary results in that paper's appendix based on satellite data yield estimated elasticities between 5% and 10% for PM2.5.

Estimates when using the PM2.5 exposure variable in levels indicate that a doubling of population density will increase particulate matter concentration by $0.73\ \mu g/m^3$ ($1.047 \times \ln(2)$). This is a substantial effect as the cross-sectional standard deviation in density in our sample is 2.2 ($\mu g/m^3$).[17] Overall, these results show that denser cities have worse environments than more sprawled out cities in terms of PM2.5 concentration.

Finally, we estimate the effect of density on the fraction of an urban area's population that lives in locations that do not comply with the EPA's NAAQS for PM2.5. We calculate this fraction for every city by aggregating from our detailed spatial data on population and PM2.5 concentration, and replace this variable a continuous outcome in Eq. (4). The resulting estimates are provided in Appendix Table A4. We obtain OLS and IV estimates of 0.4% and 1.3%, respectively, in our preferred specifications. Focusing on the IV estimates, these indicate that a doubling of density would result of an increase of roughly 1 percentage point in the fraction of population leaving in areas with non-compliant concentration levels.

16 Results provided only for the set of CBSAs that could be matched to the WLUR city definitions.
17 Results provided only for the set of CBSAs that could be matched to the WLUR city definitions.

Table 3
Between-city – 2SLS estimates.

	Log(PM2.5) – elasticities		
Log(Pop. Dens.)	0.149***	0.158***	0.137***
	(0.034)	(0.023)	(0.026)
Observations	933	933	933
	PM2.5 (exposure)		
Log(Pop. Dens.)	1.239***	1.235***	1.062***
	(0.250)	(0.167)	(0.186)
Full controls	Yes	No	Yes
State-FE	No	Yes	Yes
F-Stat	21	19	20
Obs.	933	933	933

Notes: Reports IV estimates of the effects of log density on PM2.5 exposure. The unit of analysis is the city (CBSA). For the first row of estimates, the dependent variable is the natural logarithm of PM2.5 exposure as defined in the text. The dependent variable for the second row of estimates is this variable in levels. All specifications include latitude, longitude and average maximum and minimum temperatures as controls. The specifications in columns 2 and 3 add state effects. The specification in column 3 adds a set of additional controls as detailed in the text. F-statistics for joint significance of the geological instruments in the first-stage reported in the table foot. Robust standard errors in parentheses.
***p < 0.01, **p < 0.05, *p < 0.1

4.2. Robustness

In this section, we provide a series of additional estimates to highlight the robustness of our between-city results to (i) the sets of selected geological instruments, (ii) using a first difference estimation based on a two-period panel covering the years 2000 and 2010, (iii) using an alternative IV strategy based on a historical density instrument, (iv) using monitoring station data to measure city level pollution, and (v) using population-weighted densities as our main variable of interest.

We begin with Appendix Table A5, which provides alternative estimates obtained by sequentially excluding one of our geological instruments for density. All reported estimates are pollution-density elasticities and our preferred specification continues to include state effects and the full set of controls (column 3). For the first row of estimates only aquifer presence and soil drainage capacity are included as instruments. In the second row, only earthquake risk and aquifer presence are used as instruments. Finally, in the third row only drainage capacity and earthquake risk are used as instruments. Estimates for all instrument pairs are positive and significant, as expected. Moreover, estimates in the first and third row of Table A5 are very close to those reported in Table 3. We also conduct an overidentification test to evaluate whether our instruments identify the same parameter of interest. We cannot reject the null of exogeneity, which lends further support to the validity of our geological instruments.[18]

Next, we turn to a complementary research strategy using a two-period panel for the years 2000 and 2010. The CBSA definitions used in the cross-sectional analysis above were created in 2008, so no appropriate boundaries are available for the 2000 census. We therefore need an alternative definition of the urban area, that does incorporate changes in the extent of urban areas over this decade. For this purpose, we use the definition of Commuting Zones (CZ) described in Fowler et al. (2016), which draws on commuting data from the census and the American Community Survey (ACS) to delineate these zones. We use this source because it provides methodologically consistent delineations for 2000 and 2010 and incorporates changes in the surface of urban areas.[19] Using our CZ panel, we estimate the following regression equation:

$$Ln(PM2.5^{exp}_{ct}) = \beta Ln(Density_{ct}) + \gamma X_{ct} + \eta_s \times \delta_t + \alpha_c + \varepsilon_{ct} \tag{5}$$

where c is now an index for commuting zones, α_c is a CZ fixed effect, and X_{ct} includes our basic set of controls (geographic coordinates, and average minimum and maximum temperatures) interacted with a time dummy. In our preferred specification, we also include the term $\eta_s \times \delta_t$ capturing state-specific trends in density and pollution concentration. Panel A of Table 4 reports our elasticity estimates which are positive and significant, taking a value of 0.1 in our preferred specification, roughly 4/5 of our IV estimates above. The comparison between estimates is complicated by the fact that both the nature of the variation – longitudinal vs. cross-sectional – and the units of analysis are different in both exercises. That being said, we interpret the panel results as providing further evidence of the positive link between particulate concentration and urban density. We continue to put our emphasis on our IV estimates as it is unlikely that longitudinal changes in density are exogenous in Eq. (5).

We can obtain alternative estimates of the effect of density on PM2.5 concentration using the historical density IV instead of our geologic variables.[20] Panel B in Table 4 reports the 2SLS estimates using this strategy. We continue to observe positive and

[18] When estimating the model with state effects and controls the p-value of the test is 0.25.

[19] We match 2010 with 2000 commuting zones by taking the 2000 CZ with the closest population-weighted centroids. To ensure the match indeed captures the same city, we limit the distance between centroids to under 20 km. This limits our sample to 460 commuting zones.

[20] Because some of the current United States counties were not covered in the 1880 census, the number of observations is restricted to 920 out of our original 933 CBSAs.

Table 4
Robustness - panel and historical IV.

	Log(PM2.5) - Elasticities		
A. Panel estimates			
Log(Pop. Dens.)	0.121***	0.075**	0.105***
	(0.009)	(0.033)	(0.034)
Comm. Zone FE	No	Yes	Yes
Year	Yes	Yes	Yes
State-Year FE	No	No	Yes
Obs.	920	920	920
B. Historical instrument			
Log(Pop. Dens.)	0.383***	0.257***	0.248***
	(0.045)	(0.032)	(0.029)
Full controls	Yes	No	Yes
State-FE	No	Yes	Yes
F-stat (Historical instruments)	55	39	46
Obs.	920	920	920
C. Monitoring stations			
Log(Pop. Dens.)	0.298***	0.253***	0.149***
	(0.052)	(0.049)	(0.036)
Full controls	Yes	No	Yes
State-FE	No	Yes	Yes
F-stat	49	49	52
Obs.	542	542	542

Notes: Dependent variable is the natural logarithm of PM2.5 exposure as defined in the text in all specifications. In Panel A, we report panel estimates of the density-pollution elasticity. Sample is based on a two-period panel using the time-varying definition of commuting zones in Fowler et al. (2016) for 2000 and 2010. All columns include year effects. Columns 2 and 3 include CZ effects and column 3 includes state-year interactions. In Panels B we estimate the elasticity of interest using the logarithm of historical population density (in 1880) as an instrument for density in 2010. In panel C we report IV estimates obtained using monitoring station data on PM2.5 concentration and (log) 1880 density as the instrument for density. All specifications in panels B and C include latitude, longitude and average maximum and minimum temperatures as controls. The specifications in columns 2 and 3 add state effects. The specifications in column 3 add a set of additional controls, as detailed in the text. Robust standard errors in parentheses. ***$p< 0.01$, **$p < 0.05$, *$p < 0.1$.

significant elasticities throughout. When using the historical instrument, however, the elasticity is larger than before, reaching 24% in our preferred specification. This is almost twice the size of the estimate obtained using our geological variables. We interpret this coefficient with care, given that we expect historically dense cities can have urban features such as older central heating systems which could affect air pollution directly, thus breaking the exclusion restriction. With that caveat in mind, it is still reassuring that the qualitative findings using our geological instruments are the same as those obtained using this alternative instrumentation strategy.

Our main findings are also robust to using monitoring station data to measure PM2.5. We report density effects using average concentrations obtained from the US EPA monitoring station network in panel C of Table 4. In this case we use our historic instrument for density as the geological instruments are weak in the restricted sample. The resulting elasticities are similar to those reported in Table 3, particularly after we include state effects. As stated above, monitoring station data is only available for a selected sample of typically dense locations for certain cities, hence they give a selected measurement of local pollution levels. The advantage of monitoring station data lies in that point measurements are arguably more precise than those obtained from satellites and do not rely on model-based adjustments as in Van Donkelaar et al. (2015).

Finally, we consider an alternative definition of our density variable, by computing population-weighted population densities in each city. Population-weighted densities are used in Glaeser and Kahn (2004) and Rappaport (2008). In our case, these result from taking a weighted-average of the grid cell level population densities within a city, with weights given by the fraction of the population of each grid cell. This variable can be interpreted as an estimate of the average population density faced by residents in each city (Glaeser and Kahn, 2004). We report estimates of a modified version of Eq. (2) where we use the logarithm of this variable as our measure of density. OLS and IV estimates of the resulting elasticities are provided in Appendix Table A5.[21] The results show positive and significant elasticities in all specifications, indicating that our main qualitative result does not depend on how we measure density. The elasticities are in this case almost five times as large as those using the conventional density variable, which is consistent with the main estimates being driven by high concentrations of polluters within cities.

[21] Both the aquifer and soil drainage instruments have insignificant effects on this alternative measure of density. This is not surprising given that these variables would affect the extensive rather than intensive margins of urban development – they affect where building is possible/desirable rather than the intensity of that development. IV estimates in Appendix Table A6 are obtained using our historical and earthquake instruments only.

Table 5
Within-City – 2SLS estimates.

	Log(PM2.5) – elasticities			
Log(Pop. Dens.)	0.147***	0.083***	0.317***	0.213***
	(0.055)	(0.022)	(0.047)	(0.042)
Observations	4 306 842	4 306 842	4 306 842	4 306 842
	PM2.5			
Log(Pop. Dens.)	0.694***	0.520***	1.089***	0.671***
	(0.210)	(0.115)	(0.152)	(0.126)
Full Controls	No	Yes	No	Yes
State-FE	Yes	Yes	No	No
City-FE	No	No	Yes	Yes
F-Stat	12	25	15	10
Obs.	4 325 515	4 325 515	4 325 515	4 325 515

Notes: Estimates from grid-cell level 2SLS specifications. Dependent variable in the first row of estimates is the natural logarithm of PM2.5 concentration. Dependent variable in the second row of estimates is PM2.5 concentration. All specifications include latitude, longitude and average maximum and minimum temperatures as controls. Columns 2 and 3 include state fixed effects. Columns 3 and 4 include CBSA fixed effects. The specifications in columns 2 and 4 add a set of additional controls as detailed in the text. Standard errors clustered at the city level in parenthesis. ***p< 0.01, **p < 0.05, *p < 0.1.

4.3. Within-city estimates

We complement our main between-city analysis by estimating how local changes in density affect pollutant concentration within cities. Baseline estimates of Eq. (2) obtained using OLS are reported in Appendix Table A7. The coefficient is positive and significant across specifications, indicating an elasticity of 3.6% when including both CBSA fixed effects and the full set of controls. This implies that doubling population density in a grid cell leads to a 2.5% increase in PM2.5 concentration within a city.[22]

Turning to our instrumental variable results, we first provide estimates for the first-stage coefficients in Appendix Table A8. The F-statistics for a joint significance test of the three coefficients indicates that our instruments are not weak, which has two important implications. First, it demonstrates that the relevance condition is satisfied. Second, it also supports the empirical strategy used in the between-city analysis. The logic behind the use of these instruments is based on their impact at the micro-level. For instance, when we say aquifers affect density by reducing the need to connect to municipal water networks and allowing for sprawl, we are using a within-city rationale, even if the instrument is used in a between-city analysis. The relevance of our instruments within-city is reassuring because it clarifies why they are relevant across cities.

The IV estimates for our within-city analysis are reported in Table 5. The elasticity estimates, reported in the first row, are substantially larger than those obtained under OLS, indicating a bias towards zero in our baseline estimates. This bias is again consistent with reverse causality, with pollution levels affecting the distribution of population within cities.[23] In our preferred specification in column 4, we find an elasticity of roughly 0.2, indicating that a 1% increase in population density in a grid cell increases PM2.5 concentration by 0.2%. This elasticity is sensibly larger than the between-city estimate of 0.14, indicating that the local effect of density on pollution is larger than the aggregate city-wide effect.

5. Discussion

We have shown above that dense urban development results in worse air quality. In this section, we discuss the possible underlying channels behind our reduced-form findings. We posit that four main channels can theoretically explain our results: (1) population scale-effects, (2) the composition of economic activity, (3) differences in transportation, and (4) spatial concentration of polluters in denser areas. We explore each of these possible channels below.

We begin by exploring whether our findings arise because denser cities are simply larger. If city-wide effects lead larger cities to have higher pollutant concentrations, then this could operate as a mechanism linking density to air quality. We consider two specifications to control flexibly for the total CBSA population when estimating the relationship between density and air quality. In this way, we hope to purge any density-induced changes in the total population, as well as any remaining confounders related to city size.

We start by including a fourth-degree polynomial in population in our specification and re-estimate the density — concentration elasticity. Results are provided in the top panel of Table 6. We observe that the elasticity of interest is approximately 20 percent larger after controlling flexibly for population, and not statistically different from the point estimate reported in Table 3. These results suggest scale-effects are not driving our results. We explore this possible channel further by including the logarithm of the

[22] Note that if $\beta = 0.036$, then $2^{\beta} - 1 = 0.0252$.

[23] Recent evidence on population sorting within cities in response to pollution can be found in Heblich et al. (2021).

Table 6
Controlling flexibly for population.

	Ln(PM2.5) - elasticity		
Control for polynomial in population	0.257***	0.205***	0.167***
	(0.063)	(0.040)	(0.041)
F-Stat	14	12	12
Obs.	933	933	933
	Ln(PM2.5) - elasticity		
Instrument for Ln(Population)	0.745***	0.440***	0.369***
	(0.130)	(0.141)	(0.090)
Full controls	No	No	Yes
State-FE	No	Yes	Yes
F-Stat 1	19	22	23
F-Stat 2	51	42	10
Obs.	920	920	920

Notes: 2SLS estimates obtained by modifying our main between-city specification. The dependent variable is the natural logarithm of population-weighted PM2.5 concentration as defined in the text. In the specifications reported in the top panel, we control for a 4th degree polynomial in CBSA population. To obtain estimates in the bottom panel, we include the logarithm of CBSA population as an instrumented variable and add in the log of 1880 population density as an instrument. All specifications include latitude, longitude and average maximum and minimum temperatures as controls. The specifications in columns 2 and 3 add state effects. The specification in column 3 adds the fullset of controls (listed in Section 4.1). F-statistics from first-stages, reported in the table foot. Robust standard errors in parentheses.
***p < 0.01, **p < 0.05, *p < 0.1.

population as an instrumented variable in a specification using both, geological and historical instruments. The resulting elasticities are reported in the bottom panel of Table 6. Again, we find that the coefficient of interest is larger than the one obtained when using all instruments but excluding the logarithm of the population. From these exercises, we conclude that our results are not driven by population-scale effects.

As highlighted above, a second possible mechanism linking density to pollution concentration could exist if the sectoral composition of different cities varies with their density. A substantial amount of PM2.5 pollutant emissions is produced by manufacturing and other industrial activities. If agglomeration forces for these industries are relatively more pronounced than in other sectors, then this could explain the relationship between density and PM2.5 concentration. To explore this possibility, we test whether observed differences in sectoral composition can be explained by density. We conduct three different exercises for this purpose. First, we compute the fraction of total employment devoted to manufacturing in each CBSA, by aggregating data from the County Business Patterns dataset for 2010. We substitute this variable as the outcome variable in Eq. (4), and estimate the effect of density on this measure of industrial composition by using our geological IVs. Results are provided in panel A of Table 7. In all three columns, we observe small and insignificant effects of population density on the fraction of manufacturing employment.

Clearly, not all manufacturing activities are the same, and it is still possible that changes in composition within the manufacturing sector led to differences in PM2.5 concentration across cities. To explore this possibility, we use data on industrial composition from CBP, combined with data on PM2.5 emission intensities by International Standard Industrial Classification (ISIC) industry obtained from Shapiro and Walker (2018). We then compute, for every city, the variable

$$Ln(CompPM2.5_c) = Ln\left(\sum_{i=1}^{N_I} \frac{emp_{ic}}{\sum_{j=1}^{N_I} emp_{jc}} IntPM2.5_i\right)$$

where $IntPM2.5_i$ is the intensity measure obtained from Shapiro and Walker (2018) and $\frac{emp_{ic}}{\sum_{j=1}^{N_I} emp_{jc}}$ is the fraction of employment from city c dedicated to industry i. The variable will take relatively high values in cities that specialize in industries producing large quantities of PM2.5 pollutant emissions. We replace $Ln(CompPM2.5_c)$ as the outcome in Eq. (4) and estimate the effect of density on this variable. The resulting elasticities are reported in panel B of Table 7 and become statistically insignificant when including state effects and the full set of controls. We conclude that the effect of density on the local intensity of PM2.5 polluting industries is unlikely to explain our headline results.[24]

To test the robustness of our findings for polluting intensities, we also construct an alternative measure of industrial-composition emission intensity based on the PM10 intensities reported in Levinson (2009).[25] Using these intensity measures we compute

[24] An alternative specification of our main estimating equation which includes $Ln(CompPM2.5_c)$ as a control leads to essentially the same density-concentration elasticities as those reported in Table 3 (available upon request).

[25] These are based on the World Bank's Industrial Pollution Projection System (IPPS) which reports emission intensities for 4 level 1987 SIC codes. We convert these into 2-digit NAICS 2007 intensities using the crosswalk between 1987 SIC codes and 2002 NAICS codes, combined with the crosswalk between the 2002 and 2007 NAICS codes.

Table 7
Density and the composition of polluting economic activities.

A. Employment composition			
	Fraction of manufacturing employment		
	(1)	(2)	(3)
Log(Pop. Dens.)	0.019	0.006	0.022
	(0.020)	(0.019)	(0.020)
B. Composition of polluters (PM2.5)			
	Log(PM2.5 intensity)		
Log(Pop. Dens.)	0.174*	0.026	0.072
	(0.102)	(0.092)	(0.097)
C. Composition of Polluters (PM10)			
	Log(PM10 Intensity)		
Log(Pop. Dens.)	0.192*	0.033	0.058
	(0.107)	(0.098)	(0.107)
Full controls	No	No	Yes
State-FE	No	Yes	Yes
Obs.	933	933	933

Notes: Panel A reports 2SLS estimates of the effect of density on the fraction of CBSA employment working in manufacturing. Panel B reports 2SLS estimates of the effect of density on the PM2.5 pollution intensity composition at the city level as derived from Shapiro and Walker (2018). Panel C reports estimates of the effect of density on PM10 pollution intensity composition as derived from Levinson (2009). All specifications include controls for temperature and geographic coordinates. Specification in column 2 adds state effects. Specification in column 3 adds extra set of controls. Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

$Ln(CompPM10_c)$ calculated as above, and study how this variable is affected by density. Results are provided in Panel C of Table 7. As with PM2.5 intensity, the resulting elasticity becomes small and statistically insignificant when we include state effects in the specification. To sum up, the coefficients in Table 7 indicate that potential differences in industrial composition resulting from differences in densities across cities cannot explain the reported effect of density on PM2.5 pollutant exposure.

Emissions from transport, in particular from commuter flows, can also have an impact on local pollutant concentration. Previous research indicates there is less driving in denser areas (see for example Duranton and Turner (2018) and Stevens (2017). While the relationship between density and driving-related emissions is complicated by the role of road-congestion and emissions, but current estimates of total effects do indicate a negative association (Glaeser and Kahn, 2010). Changes in transport-mode – e.g., more widespread use of mass transit in urban areas – would reinforce this effect. As a result, it is unlikely that changes in driving can explain our findings.

The mechanisms discussed above are unlikely to explain the results of our between-city analysis. However, there is an additional potential mechanism that relates to the spatial diffusion of pollution in dense environments. A higher spatial concentration of activity can translate into higher pollutant concentrations in the air because polluters are close to each other. While we cannot test for this mechanism directly without data on emissions, we do want to emphasize that this hypothesis is consistent with the large and significant positive estimates reported in our within-city analysis.

6. Health implications and costs

In this section, we assess the mortality impacts and economic costs of air pollution-induced by density, based on our estimates. Our analytical strategy is very similar to the approach taken by the US EPA in their Regulatory Impact Analysis (Environmental Protection Agency, 2012) and consists of the following two steps. First, we relate changes in pollution concentrations due to changes in population density with mortality using Concentration–Response functions (C–R functions). Second, we estimate the associated economic costs by multiplying the mortality effect by the Value of Statistical Life (VSL). C-R functions link pollution exposure (PM2.5) to mortality incidence rate (y) and are most commonly estimated using a log-linear form as follow:

$$y = B \times e^{\beta \times PM2.5} \Rightarrow \quad ln(y) = \alpha + \beta \times PM2.5$$

where $ln(y)$ is the natural logarithm of y, $\alpha = ln(B)$, β is the coefficient of interest which measures the estimated average effect of PM2.5, and B is the incidence rate of y when $PM2.5 = 0$.[26] Defining y_0 as the baseline mortality incidence rate, we can write the relationship between changes in PM (ΔPM) and mortality incidence rate (Δy) as:

$$\Delta y = y_0 - y_1 = B(e^{\beta PM2.5_0} - e^{\beta PM2.5_c})$$
$$\Delta y = \beta \times e^{\beta PM2.5_0}\left(1 - e^{-\beta(PM2.5_0 - PM2.5_c)}\right) = y_0\left(1 - \frac{1}{exp(\beta \times \Delta PM2.5)}\right)$$

[26] B can also be interpreted as a vector of covariates which may affect mortality and defined as: $B = B_0 \times e^{\beta_1 x_1 + \dots + \beta_n x_n}$ where B_0 is the incidence of y when all covariates in the model are zero, and $x_1, \dots, x_n$ are other covariates.

Multiplying the mortality incident rate by the relevant population yields the change in the incidence of mortality which is our prime objective.[27]

We follow Fowlie et al. (2019) and rely upon two influential studies that estimated mortality Relative Risks (RR) in the US. The first paper is a follow-up examination of the Harvard Six Cities study by Lepeule et al. (2012) which documents a significant statistical association between PM2.5 and mortality. Using a Cox proportional hazards model the authors found an RR of 1.14 (CI 95% [1.07,1.22]), implying that a $10\ \mu g/m^3$ annual increase in PM2.5 is associated with a 14% increased risk of all-cause mortality. The second paper by Krewski et al. (2009) is a large cohort study that used a random-effects Cox model to estimate the C–R function among the US population. The authors found a mortality RR of 1.06 (CI 95% [1.04,1.08]) which is smaller than in Lepeule et al. (2012) but still highly significant.

Using the above C-R functions and our estimates from Table 3, we analyze what would be the impact of doubling density in an average US county. To put this in perspective, this is the equivalent of changing population density in Houston to that of Chicago. Our estimates indicate this would result in increasing annual PM2.5 concentration by $0.73\ \mu g/m^3$. Our analysis suggests that the annual per capita mortality costs of doubling density, using the high and low C-R functions from Lepeule et al. (2012) and Krewski et al. (2009) in conjunction with the EPA VSL recommended estimate of 7.4 million (2006 USD), are $630 and $281, respectively. The former estimate is large and equivalent to between 17 and 39 percent of the estimated agglomeration effect on productivity for a worker earning the average wage in 2010.[28]

As prior research suggests that compact cities are linked with lower greenhouse gas emissions, we also compare our cost estimates with expected benefits resulting from reduced CO2 emission in denser cities using a back-of-the-envelope calculation. For this purpose, we build on the carbon cost-saving calculations in Ahlfeldt and Pietrostefani (2019). Using the elasticity of residential and transport energy consumption with respect to density, a US average annual per capita CO2 emissions from residential and transport energy of 25 tons and a social cost of carbon of $43, we find that doubling density leads to a cost reduction of $52.1 per capita.[29] If we restrict the costs of carbon to mortality effects only, then the benefits from doubling density amount to only $47.3, based on the upper bound estimate from Carleton et al. (2022). While these figures are only suggestive, it is worth noting that they are both substantially smaller than our estimates of the mortality costs of doubling density attributed to PM2.5. Therefore, comparing the environmental global benefits and local costs of density, our calculations indicate that the costs far outweigh the benefits.

Finally, we estimate the annual mortality costs of doubling density for each CBSA in the US separately. To do that, we use the Benefits Mapping and Analysis Program (BenMAP) which is typically used by the US EPA in its Regulatory Impact Analysis. The program is based on the same methodology explained above (including the C-R function from Lepeule et al., 2012) but also accounts for the variation in the age structures and pollution levels across cities. The results are displayed in Appendix FigureA6. As we can see from the map, the largest costs of increasing density in terms of pollution-induced mortality occur in the largest US cities.[30]

7. Conclusions

Air pollution is one of the typical congestion forces discouraging households from moving to large, dense cities. In this paper, we report estimates of the elasticity of air pollution with respect to urban density at the city level. Using spatial data on PM2.5 pollutant concentration, we estimate this elasticity to be between 0.1 and 0.25, with a preferred estimate of 0.14. Our instrumental variable estimates indicate that a doubling of density increases PM2.5 concentration in roughly half of a standard deviation. This is a large effect because densities vary widely between urban areas in the United States. For example, the CBSA around New York City has a density that is twice as large as that around Boston and nine times as large as that for Phoenix.

Our results highlight the need to incorporate the effect on air quality when discussing suburbanization, densification policies and the environmental aspects of urban planning. We provide estimates of pollution-induced costs of density which can be used in the context of cost-benefit calculations when evaluating the desirability of these policies. With our estimates, we also highlight the important distinction between local and global pollutants, the nature of the externalities they generate, and the trade-offs involved in policies trying to address them. A usual point raised by planners, policy makers and economists, is that denser cities tend to be more environmentally friendly as they produce lower levels of carbon emissions per person. Even if this is indeed the case, it

[27] Importantly, since most epidemiological studies report the relative risk (RR) for a given ΔPM and not β, we convert RR into β by using the fact that RR is simply the ratio of the two risks which yield the following relationship: $\beta = ln(RR)/\Delta PM$

[28] According to Combes and Gobillon (2015), studies on the static benefits of agglomeration economies on productivity yield an estimate range between 0.04 and 0.05 when using an empirical strategy similar to the one used in our analysis. Taking the mid-point of that range – and allowing for some extrapolation – doubling density would result in an increase of $0.045 \times ln(2)$ for average wages. Individual average wages in the US in 2010 were 52,384 USD, which yields an approximate figure of $1633 difference resulting from a doubling in population. Reported ratios result from estimating our mortality effects and dividing by this figure.

[29] To obtain this cost per capita estimate, we multiply the elasticity of 0.07 (from Ahlfeldt and Pietrostefani, 2019) times ln(2) (doubling density), times 25 (average CO2 in tones of residential and transport emission estimate from Glaeser and Kahn (2010), times the social cost of carbon.

[30] It is important to highlight that the mortality costs of density that we present here represent only part of the total cost to human health and wellbeing. Air pollution is also adversely linked with other health and economic outcomes (such as work loss days and hospital admission) which are costly. However, we have decided to focus on mortality as this is by far the costliest consequence of air pollution. To illustrate this point, Appendix Figure A7 shows the costs in terms of Work Loss Days based on C–R function from Ostro (1987) and county-specific median wages. As evident in the map, the costs are significantly lower than the mortality estimates but the spatial pattern is the same. We have also investigated several other health costs (including emergency room visits and hospital admission for cardiovascular and respiratory cases) and the costs are also very small in comparison to our mortality cost estimates.

does not necessarily mean that denser cities provide a better environment for their inhabitants. We have shown that air pollution exposure is actually higher in denser cities, indicating that there could be a trade-off between reducing a city's carbon footprint and preserving the environmental quality within the city.

Finally, a large and growing literature has provided overwhelming evidence on the adverse effects of air pollution on human health and wellbeing. In contrast, this paper studies the determinants of air pollution itself. While the former literature is necessary to understand the magnitude of the problem, studies such as ours are necessary to evaluate suitable solutions.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeem.2022.102767.

References

Ahlfeldt, Gabriel M., Pietrostefani, Elisabetta, 2019. The economic effects of density: A synthesis. J. Urban Econ. 111, 93–107.
Albouy, David, Stuart, Brian, 2014. Urban Quantities and Amenities. NBER Working Paper, (19919).
Bertaud, Alain, Malpezzi, Stephen, 2003. The Spatial Distribution of Population in 48 World Cities: Implications for Economies in Transition, Vol. 32. Center for Urban Land Economics Research, University of Wisconsin, pp. 54–55, (1).
Boarnet, Marlon G., Wang, Xize, 2019. Urban spatial structure and the potential for vehicle miles traveled reduction: The effects of accessibility to jobs within and beyond employment sub-centers. Ann. Reg. Sci. 62 (2), 381–404.
Bondy, Malvina, Roth, Sefi, Sager, Lutz, 2020. Crime is in the air: The contemporaneous relationship between air pollution and crime. J. Assoc. Environ. Resour. Econ. 7 (3), 555–585.
Borck, Rainald, Brueckner, Jan K., 2018. Optimal energy taxation in cities. J. Assoc. Environ. Resour. Econ. 5 (2), 481–516.
Borck, Rainald, Pflüger, Michael, 2019. Green cities? Urbanization, trade, and the environment. J. Reg. Sci. 59 (4), 743–766.
Borck, Rainald, Schrauth, Philipp, 2021. Population density and urban air quality. Reg. Sci. Urban Econ. 86, 103596.
Borck, Rainald, Tabuchi, Takatoshi, 2019. Pollution and city size: can cities be too small? J. Econ. Geogr. 19 (5), 995–1020.
Burchfield, Marcy, Overman, Henry G, Puga, Diego, Turner, Matthew A, 2006. Causes of sprawl: A portrait from space. Q. J. Econ. 121 (2), 587–633.
Carantino, Benjamin, Lafourcade, Miren, et al., 2018. The carbon'carprint'of suburbanization: New evidence from French cities.
Carleton, Tamma, Jina, Amir, Delgado, Michael, Greenstone, Michael, Houser, Trevor, Hsiang, Solomon, Hultgren, Andrew, Kopp, Robert E, McCusker, Kelly E, Nath, Ishan, et al., 2022. Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. Q. J. Econ. 137 (4), 2037–2105.
Castells-Quintana, David, Dienesch, Elisa, Krause, Melanie, 2021. Air pollution in an urban world: A global view on density, cities and emissions. Ecol. Econom. 189, 107153.
Chay, Kenneth Y., Greenstone, Michael, 2003. The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. Q. J. Econ. 118 (3), 1121–1167.
Chen, Shuai, Oliva, Paulina, Zhang, Peng, 2022. The effect of air pollution on migration: evidence from China. J. Dev. Econ. 156, 102833.
Chen, Jiandong, Wang, Bo, Huang, Shuo, Song, Malin, 2020. The influence of increased population density in China on air pollution. Sci. Total Environ. 735, 139456.
Ciccone, Antonio, Hall, Robert E., 1996. Productivity and the density of economic activity. Am. Econ. Rev. 86 (1), 54–70.
Cirilli, Andrea, Veneri, Paolo, 2014. Spatial structure and carbon dioxide (CO2) emissions due to commuting: An analysis of Italian urban areas. Reg. Stud. 48 (12), 1993–2005.
Clark, Lara P., Millet, Dylan B., Marshall, Julian D., 2011. Air quality and urban form in US urban areas: Evidence from regulatory monitors. Environ. Sci. Technol. 45 (16), 7028–7035.
Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, 2008. Spatial wage disparities: Sorting matters!. J. Urban Econ. 63 (2), 723–742.
Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, 2019. The costs of agglomeration: House and land prices in French cities. Rev. Econom. Stud. 86 (4), 1556–1589.
Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Roux, Sébastien, 2010. Estimating agglomeration economies with history, geology, and worker effects. In: Agglomeration Economics. University of Chicago Press, pp. 15–66.
Combes, Pierre-Philippe, Gobillon, Laurent, 2015. The empirics of agglomeration economies. In: Handbook of Regional and Urban Economics, Vol. 5. Elsevier, pp. 247–348.
de Thé, Camille Blaudin, Carantino, Benjamin, Lafourcade, Miren, 2021. The carbon ?carprint?of urbanization: New evidence from French cities. Reg. Sci. Urban Econ. 89, 103693.
Denant-Boemont, Laurent, Gaigné, Carl, Gaté, Romain, 2018. Urban spatial structure, transport-related emissions and welfare. J. Environ. Econ. Manage. 89, 29–45.
Dockery, Douglas W, Pope, C Arden, Xu, Xiping, Spengler, John D, Ware, James H, Fay, Martha E, Ferris, Jr., Benjamin G, Speizer, Frank E, 1993. An association between air pollution and mortality in six US cities. N. Engl. J. Med. 329 (24), 1753–1759.
Duranton, Gilles, Turner, Matthew A., 2018. Urban form and driving: Evidence from US cities. J. Urban Econ. 108, 170–191.
Ebenstein, Avraham, Lavy, Victor, Roth, Sefi, 2016. The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. Am. Econ. J.: Appl. Econ. 8 (4), 36–65.
Environmental Protection Agency, 2012. Regulatory impact analysis for the final revisions to the national ambient air quality standards for particulate matter.
Environmental Protection Agency, 2022. US environmental protection agency website. https://www.epa.gov/pm-pollution/particulate-matter-pm-basics, Accessed: September 2022.
Fowler, Christopher S., Rhubart, Danielle C., Jensen, Leif, 2016. Reassessing and revising commuting zones for 2010: History, assessment, and updates for US ?labor-sheds? 1990–2010. Popul. Res. Policy Rev. 35 (2), 263–286.
Fowlie, Meredith, Rubin, Edward, Walker, Reed, 2019. Bringing satellite-based air quality estimates down to earth. In: AEA Papers and Proceedings, Vol. 109. pp. 283–288.
Gaigné, Carl, Riou, Stéphane, Thisse, Jacques-François, 2012. Are compact cities environmentally friendly? J. Urban Econ. 72 (2–3), 123–136.
Glaeser, Edward L., Kahn, Matthew E., 2004. Sprawl and urban growth. In: Handbook of Regional and Urban Economics, Vol. 4. Elsevier, pp. 2481–2527.
Glaeser, Edward L., Kahn, Matthew E., 2010. The greenness of cities: Carbon dioxide emissions and urban development. J. Urban Econ. 67 (3), 404–418.
Glaeser, Edward L., Kolko, Jed, Saiz, Albert, 2001. Consumer city. J. Econ. Geogr. 1 (1), 27–50.
Glaeser, Edward L., Sacerdote, Bruce, 1999. Why is there more crime in cities? J. Polit. Econ. 107 (S6), S225–S258.
Graff Zivin, Joshua, Neidell, Matthew, 2012. The impact of pollution on worker productivity. Amer. Econ. Rev. 102 (7), 3652–3673.

Gyourko, Joseph, Saiz, Albert, Summers, Anita, 2008. A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. Urban Stud. 45 (3), 693–729.

Heblich, Stephan, Trew, Alex, Zylberberg, Yanos, 2021. East-side story: Historical pollution and persistent neighborhood sorting. J. Polit. Econ. 129 (5), 1508–1552.

Ihlanfeldt, Keith, 2020. Vehicle miles traveled and the built environment. J. Transp. Land Use 13 (1), 23–48.

Khanna, Gaurav, Liang, Wenquan, Mobarak, Ahmed Mushfiq, Song, Ran, 2021. The Productivity Consequences of Pollution-Induced Migration in China. National Bureau of Economic Research.

Konrad, Christopher Peter, 2003. Effects of urban development on floods.

Krewski, Daniel, Jerrett, Michael, Burnett, Richard T, Ma, Renjun, Hughes, Edward, Shi, Yuanli, Turner, Michelle C, Pope III, C Arden, Thurston, George, Calle, Eugenia E, et al., 2009. Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality, Vol. 140. Health Effects Institute Boston, MA.

Lepeule, Johanna, Laden, Francine, Dockery, Douglas, Schwartz, Joel, 2012. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. Environ. Health Perspect. 120 (7), 965–970.

Levinson, Arik, 2009. Technology, international trade, and pollution from US manufacturing. Amer. Econ. Rev. 99 (5), 2177–2192.

McDonald, John F., 1989. Econometric studies of urban population density: A survey. J. Urban Econ. 26 (3), 361–385.

Melo, Patricia C., Graham, Daniel J., Noland, Robert B., 2009. A meta-analysis of estimates of urban agglomeration economies. Reg. Sci. Urban Econ. 39 (3), 332–342.

Ostro, Bart D., 1987. Air pollution and morbidity revisited: a specification test. J. Environ. Econ. Manage. 14 (1), 87–98.

Pope, C.A., Bates, David V., Raizenne, Mark E., 1995. Health effects of particulate air pollution: time for reassessment? Environ. Health Perspect. 103 (5), 472–480.

Rappaport, Jordan, 2008. A productivity model of city crowdedness. J. Urban Econ. 63 (2), 715–722.

Rosenthal, Stuart S., Strange, William C., 2008. The attenuation of human capital spillovers. J. Urban Econ. 64 (2), 373–389.

Sarzynski, Andrea, 2012. Bigger is not always better: a comparative analysis of cities and their air pollution impact. Urban Stud. 49 (14), 3121–3138.

Schlenker, Wolfram, Walker, W. Reed, 2016. Airports, air pollution, and contemporaneous health. Rev. Econom. Stud. 83 (2), 768–809.

Shapiro, Joseph S., Walker, Reed, 2018. Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade. Amer. Econ. Rev. 108 (12), 3814–3854.

Stevens, Mark R., 2017. Does compact development make people drive less? J. Am. Plan. Assoc. 83 (1), 7–18.

Sullivan, Daniel M., Krupnick, Alan, 2018. Using Satellite Data to Fill the Gaps in the US Air Pollution Monitoring Network. Resources for the Future Working Paper, 18.

Tschofen, Peter, Azevedo, Inês L., Muller, Nicholas Z., 2019. Fine particulate matter damages and value added in the US economy. Proc. Natl. Acad. Sci. 116 (40), 19857–19862.

United Nations, 2018. World urbanization prospects: The 2018 revision.

Van Donkelaar, Aaron, Martin, Randall V, Spurr, Robert JD, Burnett, Richard T, 2015. High-resolution satellite-derived PM2. 5 from optimal estimation and geographically weighted regression over North America. Environ. Sci. Technol. 49 (17), 10482–10491.

World Bank, 2022. The Global Health Cost of PM2. 5 Air Pollution: A Case for Action beyond 2021. The World Bank.

World Health Organization, 2022. World health organization website. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health, Accessed: September 2022.

Zheng, Siqi, Wang, Rui, Glaeser, Edward L., Kahn, Matthew E., 2011. The greenness of China: household carbon dioxide emissions and urban development. J. Econ. Geogr. 11 (5), 761–792.