Regional Science and Urban Economics 86 (2021) 103596

Contents lists available at ScienceDirect

Regional Science and Urban Economics

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

Population density and urban air quality☆

Rainald Borck [a,*], Philipp Schrauth [b]

[a] University of Potsdam, CESifo and DIW Berlin, Germany
[b] University of Potsdam, Germany

ARTICLE INFO

JEL classification:
Q53
R12

Keywords:
Population density
Air pollution

ABSTRACT

We use panel data from Germany to analyze the effect of population density on urban air pollution (nitrogen oxides, particulate matter, ozone, and an aggregate index for bad air quality [AQI]). To address unobserved heterogeneity and omitted variables, we present long difference/fixed effects estimates and instrumental variables estimates, using historical population and soil quality as instruments. Using our preferred estimates, we find that the concentration increases with density for NO_2 with an elasticity of 0.25 and particulate matter with elasticity of 0.08. The O_3 concentration decreases with density with an elasticity of −0.14. The AQI increases with density, with an elasticity of 0.11–0.13. We also present a variety of robustness tests. Overall, the paper shows that higher population density worsens local air quality.

1. Introduction

Are bigger and more densely populated cities better or worse places to live? Over the last centuries, the world has become more and more urbanized, as agglomeration benefits have drawn households to bigger cities. The urban economics literature on these agglomeration benefits is huge. Yet, in order to predict equilibrium and optimal sizes of cities, robust evidence is needed on the costs as well as the benefits of agglomeration, and much less seems to be known about the costs.[1] Kahn (2010) documents that in the US, larger cities have longer commuting times, higher pollution levels and higher crime rates. In this paper, we follow this line of research and analyze one particular element of the costs of agglomeration, namely, the effect of population density on air pollution. As we document below, there is hardly any evidence that credibly estimates the causal effect of density on pollution. We aim to fill this gap.

Air pollution is an acute phenomenon in many cities worldwide. Megacities in developing countries suffer from particularly high pollution levels. But even in developed countries, where urban air pollution has fallen over the last decades, high pollution levels keep occurring. Cities in Germany and other European Union countries have been subject to a variety of legal proceedings against transgressions of pollution thresholds. Therefore, the relation between urban structure and pollution concentration is an important policy issue.

Air quality is obviously an important determinant of city life. Air pollution causes severe health problems, most notably heart diseases, strokes, chronic obstructive pulmonary disease, lung cancer, and respiratory infections. According to the WHO, in 2010 air pollution caused 600,000 premature deaths in Europe alone and costs European economies US$ 1.575 trillion per year (WHO Regional Office for Europe and OECD, 2015). The European Environment Agency estimates that in Germany, particulate matter ($PM_{2.5}$) caused 66,000 premature deaths in 2013.[2] This shows the potential economic benefits of using policies to reduce air pollution. The first best policy would be to internalize pollution externalities, e.g. through Pigovian taxes or pollution licenses, but absent first-best prices, the effect of urban structure on pollution is obviously relevant for social welfare.

The pollutants we study are produced in a variety of industrial and non-industrial processes. Nitrogen oxides are produced in various combustion processes but are predominantly produced by traffic with a

☆ We thank Gabriel Ahlfeldt and two referees as well as participants at seminars in Potsdam, Max Planck Institute for Tax Law and Public Finance in Munich, the workshop "Cities and the environment" in Potsdam and the UEA meetings in Düsseldorf and Philadelphia for helpful comments. We are grateful to the German Research Foundation (DFG) for financial support.

* Corresponding author.
E-mail addresses: rainald.borck@uni-potsdam.de (R. Borck), schrauth@uni-potsdam.de (P. Schrauth).

[1] Ahlfeldt and Pietrostefani (2019) contain a useful synthesis of the research on benefits and costs of population density.

[2] See https://www.eea.europa.eu/themes/air/country-fact-sheets/germany.

https://doi.org/10.1016/j.regsciurbeco.2020.103596
Received 17 February 2020; Received in revised form 7 October 2020; Accepted 11 October 2020
Available online 21 October 2020

share of about 38%. Other sources are agriculture, as well as power generation plants and combustion processes in different industries. Particulate matter is produced by various industrial processes as well as burning of fossil fuels for heating or energy production.[3] Apart from combustion processes, they also arise through dispersion of dust on streets and tire wear of cars. Ground level ozone is created by chemical reactions between oxygen and nitrogen oxides (emitted for instance by cars) and volatile organic compounds (VOCs, which arise for instance in paints or in gasoline exhausts, but are also naturally emitted by trees in woods). Thus, human activity is the major source of bad air quality. Adverse health effects are the main reason to worry about pollutant exposure (World Health Organization, 2003). For particulate matter, all levels of exposure may lead to negative health effects but long-term threshold levels of 20 (PM_{10}) and 10 ($PM_{2.5}$) $\mu g/m^3$ were set by the WHO in order to significantly reduce adverse health effects. High levels of particulate matter affect the human respiratory system and lung.[4] NO_2 is a toxic gas, which is damaging to human health, especially at very high levels. Furthermore, nitrogen dioxide is a precursor for several other pollutants including ozone, which have been shown to have adverse health effects (World Health Organization et al., 2006). While older studies mainly found health effects of NO_2 on animals (World Health Organization et al., 2006), more recent studies also find significant health effects on humans (Costa et al., 2014). Ground-level ozone is linked to breathing problems, asthma, reduced lung function and respiratory diseases.

The effect of city size or population density on air quality has only recently become the subject of research in economics and other disciplines, and the findings have partly been contradictory (see the next section). In addition, much of the empirical literature uses cross-sectional data, sometimes from several countries, which thwarts the causal interpretation of estimated coefficients. In this paper, we estimate the effect of population density on ground-level pollution (NO_2, PM_{10}, $PM_{2.5}$, O_3, and an air quality index – [AQI]) for German cities, using rich panel data from 2002 to 2015. We focus on these four pollutants for several reasons. As discussed above, particulate matter and ozone have been identified to be particularly damaging to human health. Nitrogen dioxide has been the source of ongoing discussions about driving bans in Germany and other European countries.[5] It is interesting to study the effect of population density on these pollutants for two reasons: first, they may affect health differently (see above), and second, they may be differently affected by population density, which is indeed what we find (see below). Since in addition to individual pollutants, city residents and policy makers will be interested in overall air quality, we also study the effect of density on the aggregate AQI.

We start by presenting OLS estimates. However, these may be biased due to omitted variables or reverse causality, so we also estimate long difference (and fixed effects) regressions to control for unobserved heterogeneity that affects density and pollution. We also run instrumental variables (IV) regressions, using historical population density as well as soil quality as instruments for current population density (see Combes et al., 2010). According to our preferred estimates from the IV regressions, population density increases NO_2 and PM_{10} concentrations with elasticities of 0.25 for NO_2 and 0.08 for PM_{10}. The effect of density on smaller particulates ($PM_{2.5}$) is less precisely estimated, which is due to the more recent and less extensive measuring net. However, if we use satellite data to measure $PM_{2.5}$, we find a density elasticity of 0.08 for these smaller particulates as well. For O_3, we find a negative density effect with an elasticity of -0.14.[6] The AQI index for bad air quality increases with density, with an elasticity of 0.11 for background and 0.13 for traffic stations. In summary, we find that population density worsens air quality in German cities. We perform a variety of robustness checks which confirm our main findings.

The rest of the paper is structured as follows. The next section reviews the related literature. Section 3 presents some theoretical considerations on the link between population density and pollution concentration. Section 4 presents the data and estimation methods. Our regression results are shown in Section 5, and the last section concludes.

2. Related literature

We contribute to the growing literature that examines the interaction of city structure and environmental pollution. On the theoretical side, Borck and Pflüger (2019) analyze the channels through which population size affects pollution. In general, pollution may increase or decrease with population size (see also the model in the next section). Larson and Yezer (2015) simulate the implications of city size and density for energy use. They find that per-capita energy use falls when population density increases due to greenbelts or relaxed building height limits.[7]

On the empirical side, several recent contributions have looked at the relation between city size or density and pollution. Some papers have looked at household energy consumption and mostly found that residents of denser cities consume less energy per capita (Glaeser and Kahn, 2010; Blaudin de Thé et al., 2018). The reasons are that residents of densely populated cities may consume less fuel due to the availability of public transport systems and possibly shorter commutes, and use less residential energy because dwellings are smaller and high-rise apartment buildings are more energy efficient. Indeed, per capita fuel consumption and automobile utilization have been found to be significantly lower in more densely populated cities due to the availability of public transport and shorter commutes to work on average (Newman and Kenworthy, 1989; Karathodorou et al., 2010), and public transport has been found to reduce air pollution (Bauernschuster et al., 2017; Borck, 2019).

Several studies have examined the relation between city size and CO_2 emissions with conflicting results, e.g. Gudipudi et al. (2016) and Oliveira et al. (2014). Both of these papers use cross-sectional data. Borck and Tabuchi (2018) use panel data from US metropolitan areas. They find that per capita CO_2 emissions decrease with city size.

Another set of papers examines the effect of population size and other explanatory factors on air quality. Population size has been found to be positively correlated with pollution for particulates (Glaeser, 1998) and NO_X (Lamsal et al., 2013; Sarzynski, 2012). Population density, however, has been found to be negatively correlated with NO_X (Sarzynski, 2012; Ewing et al., 2003; Hilber and Palmer, 2014) and PM_{10} (Hilber and Palmer, 2014). On the other hand, Ahlfeldt and Pietrostefani (2019) and Carozzi and Roth (2019) found a positive effect of population density on particulate ($PM_{2.5}$) concentration with an elasticity of around 0.13. These papers use different samples and methods, but with the exception of Hilber and Palmer (2014) and Carozzi and Roth (2019), all are based on cross-sectional OLS estimates.

To our knowledge, the only two papers other than ours that seriously tackle causality are the unpublished paper by Carozzi and Roth (2019) and the now defunct working paper by Hilber and Palmer (2014). While Carozzi and Roth (2019) use IV (and fixed effects) estimates with geological instruments, Hilber and Palmer (2014) use fixed

[3] There are also natural sources such as volcanoes, dust storms or wildfires.

[4] See e.g. Pope III and Dockery (2006) for a summary of the health effects of particulate matter. For one among many recent studies, see Lagravinese et al. (2014).

[5] Other pollutants like SO_2 are associated primarily with industrial production and their importance has subsided over the last decades. For most other pollutants, there is no developed measuring system.

[6] While NO_2 is a chemical precursor of O_3, the preconditions for ozone formation are more favorable outside of cities, see fn. 36 below. This explains why we find that denser cities have higher NO_2 but lower O_3 concentrations.

[7] For papers that study population density and pollution in a welfare maximization framework, see Borck and Brueckner (2018) and Schindler et al. (2017).

effects regressions. Even fixed effects, however, may be biased if there are time varying omitted variables that affect density and pollution. The only paper besides ours that also uses instrumental variables is Carozzi and Roth (2019) who study the effect of population density on particulate pollution in US cities. The instruments they use – aquifers, earthquake risk, and soil drainage capacity – differ slightly from ours. Moreover, their main analysis is based on satellite data while ours is based on monitor readings. The latter presumably measure pollution more accurately and also contain other pollutants besides particulates. On the other hand, the placing of monitors may be non-random which could bias the estimates.[8] Even if the method is similar, the two papers present estimates from the US and Germany, two countries with different city systems and energy use and pollution patterns.[9] Finally, our paper contains data on other pollutants as well (PM_{10}, NO_2, and O_3 and overall air quality), so the findings of the studies can be viewed as complementary.

In summary, we think that many existing contributions to the literature have only limited value in identifying causal effects of population on pollution. In fact, in a survey of the economics of density, Ahlfeldt and Pietrostefani (2019) argue that pollution is one of the areas where more evidence on the effects of density is needed.

3. Theoretical considerations

In this section, we present results from a simple urban economic model of city structure and pollution, building on Borck and Brueckner (2018), Borck and Pflüger (2019), Borck and Tabuchi (2018) and Larson and Yezer (2015). We combine central aspects from these models, and extend them insofar as here we focus on pollutant concentration instead of emissions. More details are in Appendix A. Consider a circular open monocentric city with N workers who commute to the central business district (CBD) for work. Households have utility $v(c, q, \mathcal{P})$ over consumption, c, square meters of housing floor space, q, and pollution concentration $\mathcal{P}$ (see below). A household who lives at x km from the CBD incurs commuting costs tx and pays land rent $p(x)$. Mobility ensures that all households attain the same utility level throughout the city.

Housing is produced by profit maximizing developers using capital and land under perfect competition. They pay land rent $r(x)$ at distance x and an invariant price i per unit of capital. In equilibrium, land rent at the city border, $r(\overline{x})$ must equal the opportunity cost of land r_A. Let γ be the share of land in the city that is available for housing development at each distance x. We will use this parameter to induce a change in population density: as more land is available for housing, developers build more housing at each distance from the CBD, which leads to increasing densities.[10]

Let u be the outside utility residents can obtain in the rest of the economy. Then, population in the city adjusts through migration such that city residents obtain the same utility u everywhere. This canonical model produces a city where in the city center buildings are tall, dwellings small and population density high.

We assume that emissions equal the sum of emissions from commuting and residential energy.[11] Commuting emissions are assumed to be proportional to the sum of total commuting distances for all households, weighted by the emissions intensity of commuting; likewise, residential emissions are assumed proportional to total residential energy demand (itself assumed proportional to housing floor space), weighted by the emissions intensity of energy use. Pollution concentration in a city is the sum of total emissions divided by land area.[12]

Suppose now that we increase γ, the share of land available for housing. For instance, government might use zoning policies to increase the floor-area-ratio. As a result, more housing will be built and population density increases. Since increased housing supply reduces rents, residents' utility would rise, which induces in-migration from the outside economy to restore utility to the reservation level u. In the new equilibrium, the city boundary decreases, as in-migration is not strong enough to offset the increased housing supply. Since the city population has increased, its average density also increases. See Appendix A for details.

What happens to pollution concentration? Total emissions rise, as more housing is built which increases residential energy use. Moreover, since there are more residents, aggregate commuting rises due to in-migration. In sum, aggregate emissions increase. Furthermore, since the city's land area decreases, pollution concentration increases. Hence, the model predicts that cities with higher population density will have a higher pollutant concentration.

However, this model ignores some possible countervailing forces. First, due to their higher density, bigger cities tend to have a more extensive supply of public transit due to economies of scale and density. Since transit typically produces lower emissions per person kilometer than automobiles, this would tend to decrease traffic emissions, all else equal (Bauernschuster et al., 2017; Borck, 2019). In this vein, Blaudin de Thé et al. (2018) find that denser cities have better transit networks and lower car-related emissions. Second, in denser cities households especially in the city center tend to live in high-rise buildings that are more energy efficient than the detached single family homes that dominate in sparsely populated cities (see Borck and Brueckner, 2018). In Appendix A, we show, however, that while including these two forces in a stylized way attenuates the density-pollution relation, it does stay positive for realistic parameter values.[13]

In summary, we predict that, unless energy efficiency or public transit shares increase very strongly with density, higher population density should be associated with higher pollution concentration.

The model also highlights some issues in the estimation of the relation between pollution and density. There may be unobserved shifters of density that could be correlated with pollution. For instance, cities might have differing land use and environmental policies. If we cannot observe these policies and they are correlated with density, the estimation would be biased. Secondly, density is endogenous. As the model shows, density reacts to changes in exogenous parameters, such as agricultural land rent, but also to shocks to pollution concentration, which affects residents' utility and therefore leads to migration into or out of the city. We will address these issues in our estimation below.

The empirical literature has – largely descriptively – shown positive as well as negative correlations between density and pollution. To

[8] To mitigate the latter problem, we will include some station characteristics, such as distance to city center, station type and distance to main roads in our regressions (see below). Interestingly, Carozzi and Roth (2019) also use monitor reading data as a sensitivity check and find a slightly reduced effect of density on pollution. We also rerun our main regressions using satellite data and find results that are similar to those obtained with the monitor data, see Section 5.2.

[9] European and American cities differ along a number of dimensions. For historical reasons and different policies (for instance, planning policies, energy taxation, public transport investment), city structure, population density, commuting behavior including distances and mode shares, housing patterns and energy usage all differ between the US and Europe (Nivola, 1999; Gordon and Cox, 2012). Therefore, whether and how density affects energy use and pollution differently is an open question.

[10] More precisely, the increased availability of land leads to buildings that are lower, but since more space is available for housing, more will be built in total at each distance from the CBD.

[11] Borck and Pflüger (2019) in addition consider emissions from industrial and agricultural production, and intercity goods transport. Note also that we abstract from congestion, see e.g. Larson and Yezer (2015).

[12] This assumption is for simplicity. In reality, how emissions diffuse over space and time is obviously a more complicated process.

[13] In particular, in our parametrization, we assume that both commuting and residential emissions are constant elasticity functions of average density. As long as these elasticities are not lower than −1, we find that pollution increases with density. In their survey, Ahlfeldt and Pietrostefani (2019) suggest values of −0.07 for both.

shed more light on this question, we empirically examine the relation between density and pollution for a panel of German communities in the next sections.

4. Data and estimation

4.1. Data

We use administrative panel data from Germany for the period 2002–2015. While we have hourly data collected by monitoring stations for our pollutants of interest in Germany, our regional data, in particular population density, is available on a yearly basis for the roughly 400 German districts (Landkreise).

4.1.1. Emissions data

We obtained hourly emission data from the German Environmental Agency (Umweltbundesamt, UBA) for the years 2002–2015. These data are collected via a net of measurement stations throughout Germany for different pollutants.[14] Measurement stations are special monitors that lie either at streets and transport axes and measure pollution caused mainly by vehicles (traffic stations), or are dispersed throughout cities to record the overall level of city pollution at representative places (background stations). There are also stations close to industrial sites (industrial stations), but these are less numerous than traffic and background stations. The UBA also classifies the areas in which the stations are located into rural, urban and suburban areas, which we explicitly control for in our analysis. Pollutants taken into account in this paper are nitrogen dioxide (NO_2), particulate matter with diameter less than 10 μm (PM_{10}), particulates with diameter of less than 2.5 μm ($PM_{2.5}$), and ozone (O_3). We also follow common practice and construct an air quality index from these pollutant concentrations.

The availability of average hourly emissions enables us to control for differences in emission patterns, for example due to differences between peak and off-peak periods and workdays versus weekends. These variables are added to our regressions as indicator variables for each day of week and each hour of day. Furthermore, hourly emission data can be matched to weather information in more detail than lower frequency data, so we are better able to control for weather effects on emissions. The specific matching approach and the importance of taking into account weather variables are explained in Subsection 4.1.2 and in Appendix C.

We have an unbalanced panel of stations and keep only stations with more than two years of observations so we can apply long difference and fixed effects estimations.[15] In order to rule out the possibility that results are driven by seasonal forces that differently affect stations, we add a month dummy to our set of control variables.

We furthermore delete outliers from the sample. These are values above 500 $\mu g/m^3$ for particulate matter, which only occur if there is a large fire or another idiosyncratic source of high pollution (for instance New Year's eve fireworks).

Air pollution thresholds. In addition to pollution concentration levels, we will also look at extreme values, in particular, instances of transgressions of official thresholds.

Thresholds set by the EU have entered into force in 2005 (PM_{10}) and 2010 (NO_2). Global guidelines by the World Health Organization (WHO) were updated in 2005. Threshold values and their transgressions may be of particular interest, as they are supposed to be based on evidence on the health effects of pollution. If health effects increase non-linearly after the threshold is crossed, analyzing these transgressions is of particular interest. Even without any nonlinear health effects, insofar as the thresholds are legally binding, jurisdictions have a special interest in them since in case of transgressions they may be sued, as local and state governments in Germany and other EU countries have been recently.

The World Health Organization (WHO) has published guidelines for pollution concentration levels based on potential health threats. For PM_{10}, these are 20 $\mu g/m^3$ for the annual mean concentration and 50 $\mu g/m^3$ for the 24-h mean concentration. $PM_{2.5}$ is more aggressive to human health, so the thresholds are lower. The WHO recommends the annual mean pollution level to lie below 10 $\mu g/m^3$ and the 24-h mean to be lower than 25 $\mu g/m^3$. For NO_2, the guidelines contain an annual mean value of 40 $\mu g/m^3$ and a 1-h mean value of 200 $\mu g/m^3$.[16]

Air quality index. We calculate the annual air quality index (AQI), following Elshout et al. (2012). As is common practice, we calculated the AQI for traffic and background stations separately, as they use different subindices for their computations. For traffic stations, the index is the average of the NO_2 and PM_{10} concentrations divided by 40, and a subindex, which takes into account the number of days on which the PM_{10} concentration is above 50 $\mu g/m^3$.[17] For background stations, the index additionally contains an ozone subindex which accounts for the number of days with an 8-h average value greater than or equal to 120 $\mu g/m^3$.[18] The higher the total air quality index, the worse is the overall air quality. An index of less than or equal to one indicates compliance with EU standards on average, while an index above one indicates that air quality is worse on average than mandated by EU guidelines.[19]

4.1.2. Weather data

Ambient concentration of emissions is affected by weather conditions. As Auffhammer et al. (2013) argue, it is necessary to include all available weather variables in a regression, since weather variables are themselves correlated over time and space.[20] Particulate matter for example is literally washed away on very rainy days or blown out of the city on very windy ones. The concentration of NO_2 on the other hand depends on temperature and sunlight as it is one crucial precursor of ozone (O_3) formation, which depends on sunshine and therefore occurs mainly on hot and sunny days in summer.[21] The German Meteorological Service (DWD in German) provides data from its various weather and precipitation stations. We thus have hourly data on temperature, air pressure, rainfall, snowfall, sunshine, and wind.

[14] Below, we also compare our main outcomes to those with pollution measures obtained from satellite data. See Section 5.2.

[15] We repeated our OLS and IV regressions without this restriction, but results were not affected.

[16] The guidelines set by the EU are less strict but binding for its member states. The EU has published an annual threshold of 40 $\mu g/m^3$ and a 1-h threshold of 200 $\mu g/m^3$ for NO_2. The latter is allowed to be exceeded up to 18 times per year. For PM_{10}, there is an annual threshold value of 40 $\mu g/m^3$, while the 24-h-mean should lie below 50 $\mu g/m^3$ with an allowance of 35 exceedances annually. For $PM_{2.5}$ there is only an annual threshold of 25 $\mu g/m^3$. For a full list of air quality standards in the EU, see https://ec.europa.eu/environment/air/quality/standards.htm.

[17] The maximum number of days with average daily values above 50 $\mu g/m^3$ allowed by the EU is 35 at the moment. The value of this sub-index is $\frac{\log(\text{Nr. of days}+1)}{\log(36)}$. See Elshout et al. (2012) for a brief discussion why the subindex is calculated like this.

[18] The subindex is calculated as $\frac{\#\text{days with 8-hour average} \geq 120}{25}$, because the EU target is not to exceed 25 days a year with values above 120.

[19] Since the AQI is an annual index by district, we do not control for station characteristics or weather variables.

[20] It might be that some weather variables are themselves affected by population density, for instance, if denser cities are hotter or more or less windy. Therefore, we also ran regressions without weather controls. However, we do not find that weather changes our results, which is why all of our outcomes include weather controls (results without weather controls are available upon request).

[21] As Auffhammer and Kellogg (2011) note, ozone creation needs a certain amount of NO_2 and of other VOCs. If climatic preconditions are not given, NO_2 levels therefore stay high. Furthermore, at great heat, plants are less able to absorb ozone, which increases ozone concentration in the air on very hot days.

While Auffhammer and Kellogg (2011) and Wolff (2014) control for daily weather, we are able to match hourly weather variables with hourly emissions. The matching of emissions monitors and weather stations is described in Appendix C.

4.1.3. Other control variables

We can include various additional control variables in our regressions. An important determinant of recorded pollution concentration levels is the physical location of a monitoring station. We can control for a set of station-specific factors such as the distance to the CBD,[22] whether the station lies in an urban or a suburban area, the station type (traffic, background, or industry, see Section 4.1.1), and the distance to the closest major road (*Bundesstraße* or a street of similar size).[23]

In Germany, over the course of the past 15 years, many cities introduced low emission zones (*Umweltzonen*, LEZ hereafter). Those zones were established in order to lower high levels of particulate matter by restricting city access to cars that have particle filters.[24] Using maps, we assign to each monitoring station an indicator for whether or not it lies in a LEZ. Including the emissions zone indicator makes sense as policy schemes that differ between cities affect city level pollution. For instance, Gehrsitz (2017) and Wolff (2014) found that such badges significantly lower PM_{10}-levels (but not other pollutants) within cities after their introduction.

In order to control for economic drivers of pollution, we can control for district level GDP, unemployment rate and average private household income within a district. Moreover, we collected the vote share for the Green Party, in order to capture for the potential sorting of 'green' households into cities. We also have a measure of the number of public transit users as a share of total inhabitants per year. Another variable that we control for in the robustness section is the area of green space in a district. Green space may affect pollution in several ways. On the one hand, plants can absorb particles and thus mitigate pollution, and green space may lower temperatures which can also lower pollutant concentrations; on the other hand, plants and trees can generate VOCs, which then react with NO_2.[25] As a consequence, O_3 concentrations may rise, while NO_2 concentrations may fall.

Lastly, we can also control for the presence of coal-fired power stations in a district and the distance of a monitoring station to a coal-fired power station.[26] Since burning coal leads to high pollution levels, this might take out some variation that is caused by the presence of coal mines.

Some of our potential control variables may be 'bad controls', since they are themselves affected by density. We therefore choose to include these variables only in robustness checks (see Section 4.3 below). We can also think of some of these as mediating variables through which the effect of density works on pollution. We comment on this further in Section 5.2 below.

[22] Our main geographic units are districts, which often contain several cities or towns. Therefore, we define the CBD as the centroid of the most densely populated municipality within a district. For district free cities, the CBD is defined as the centroid of the city.

[23] To construct the distance, we use maps provided by the Federal Office of Cartography and Geodesy (©GeoBasis-DE/BKG - 2018).

[24] There are three different levels of LEZs: green, yellow and red with green being the most and red the least restrictive. Thus, these zones differ in the quality of the particle filters of cars. We have the exact dates when a red, yellow or green low emission zone was implemented.

[25] Natural sources of VOCs are e.g. broadleaf trees and conifer, which produce VOCs via evaporation.

[26] Since we do not have exact geo-coordinates of those power stations, we calculated the distance of the monitoring station to the centroid of the closest postal code region that accommodates a coal-firing power plant. Postal code regions are relatively small administrative units, with more than 8200 in our sample (compared to about 400 districts) and an average size of about 65 square kilometers.

Table 1
Descriptives.

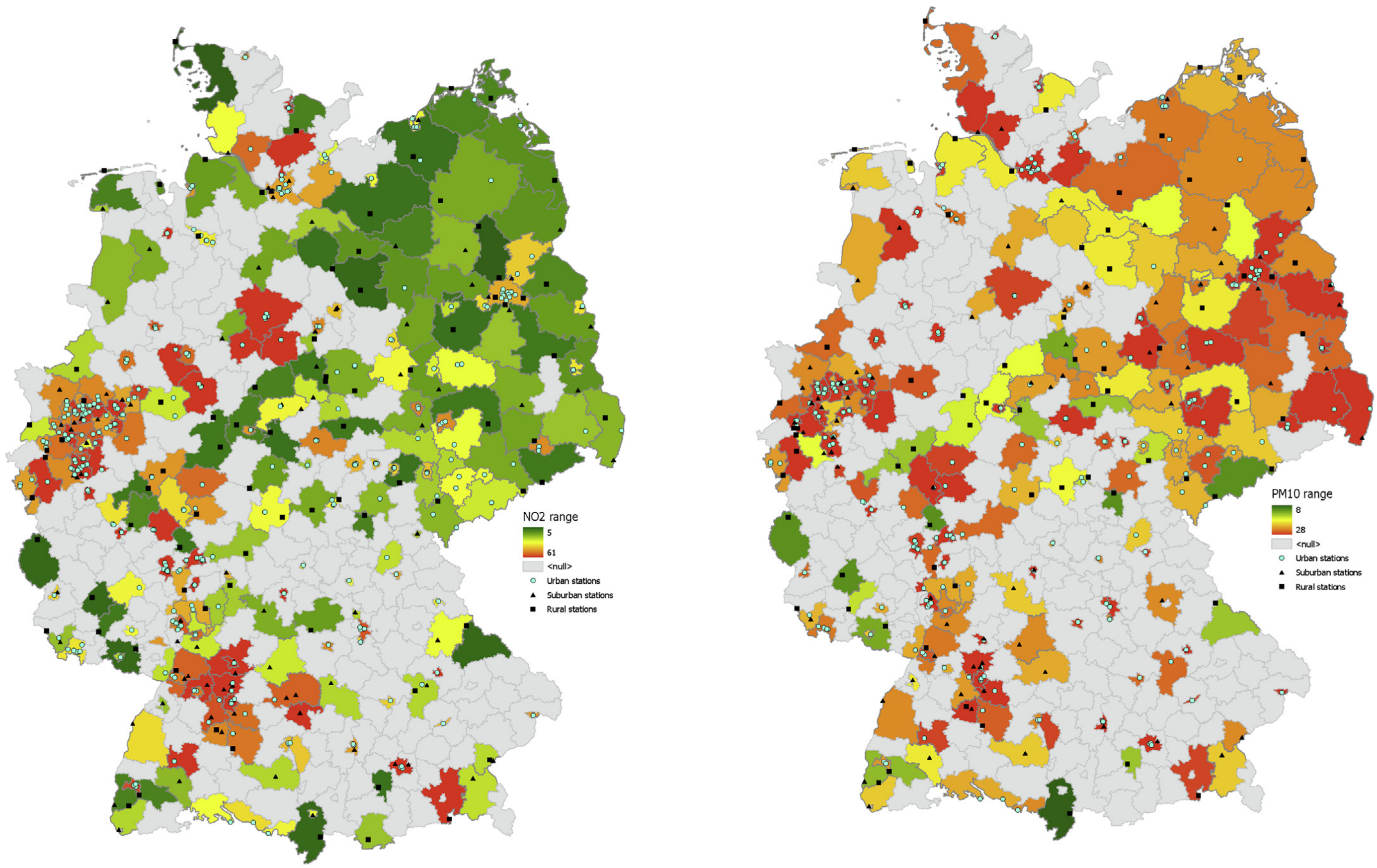
	NO2	PM10	PM2.5	O3
Overall Stations	623	533	147	409
Background	360	290	80	340
Industrial	43	45	15	26
Traffic	220	198	52	43
Districts	269	247	109	251
Urban Districts	88	85	51	72
Labor market regions	128	125	77	126
Functional urban areas	63	61	38	53
Stations in LEZ	93	94	26	34
Whole Sample				
Mean Pollution	28.42	23.20	14.83	47.24
S.D. Pollution	15.52	5.964	2.963	10.10
Mean Popdensity	2590.2	2543.6	2383.0	2249.0
S.D. popdensity	1337.6	1325.7	1322.5	1262.2

4.2. Descriptives

Table 1 presents monitoring stations in our sample and how they are distributed. The coverage of monitoring stations varies widely with NO_2 being measured by the most extensive net of monitoring stations (623), while $PM_{2.5}$ is measured by only 147 monitors, as monitoring of this pollutant only started in the mid 2000s with an extending network since then. The number of monitoring stations within the samples is also reflected in the number of districts (*Landkreise*), which are our main regional unit of analysis. In Germany, there are currently 401 districts including urban districts.[27] In addition to using districts, we also use labor market regions as defined by Kosfeld and Werner (2012) in order to check whether our results are driven by the geographical delineation of cities (there are 141 labor market regions of which we cover up to 128 in our analyses; the regions not covered do not contain a monitoring station for any pollutant). These are defined as metropolitan regions made up of several districts with large commuting flows between them (see Kosfeld and Werner (2012)). As can be seen from the table, most of these contain at least one monitoring station for NO_2, while $PM_{2.5}$ stations are only present in about half of the labor market regions.

For a first visual impression about pollution patterns and how pollution is recorded, Fig. 1 shows annual mean concentration levels of NO_2 and PM_{10} in 2015 and the distribution of monitoring stations in each district. Analogous maps for $PM_{2.5}$ and O_3 are shown in Fig. A2. The small districts in the maps are mostly urban municipalities, which are more densely populated than other parts of the country. These areas also show high concentrations of pollution. Furthermore, the historical industrial regions in West Germany and the automotive center around Stuttgart show high values of PM_{10} and NO_2. The figures also reveal pollution patterns that are clearly not related to high population densities. For instance, PM_{10} shows high concentration levels in less urbanized districts in East Germany. These high levels might be caused by the proximity to coal-fired power stations in these areas. We will control for the presence of coal fired power plants in the robustness section of the

[27] The German administrative system distinguishes between districts (*Landkreise*) and district-free cities or urban districts (*kreisfreie Städte*). The latter are entities where the 'district' consists of a single (large) city, while *Landkreise* contain several jurisdictions. The table shows the number of districts, including those urban districts, which are covered by monitoring stations. In the main analysis, we use both types, but in a sensitivity check, we also rerun our main regressions for district-free cities only.



(a) Sample of NO_2 (b) Sample of PM_{10}

Fig. 1. Monitoring stations (black dots) and mean NO_2 (left) and PM_{10} (right) concentration levels in 2015.

paper.[28]

Fig. 2 shows scatter plots for the four pollutants (logged mean pollution on district level) against logged population density. The figure shows that density is positively correlated with NO_2 and particulates, although for the latter there is more noise. On the other hand, ozone concentration is negatively correlated with population density. Below, we will analyze whether those results hold up in multivariate regressions.

4.3. Estimation

4.3.1. Basic regressions

We now turn to estimating the model. While the pollution monitor readings are hourly data, our main variable of interest, population density, is available only annually. Therefore, and in order to reduce computational burdens, we first regress hourly pollution on hourly weather data, as well as time indicators. Following Auffhammer et al. (2013), the extensive set of weather and weather-interaction variables includes the hourly level of precipitation, sunshine, wind-speed, cloudiness, air pressure, and temperature at weather stations, as well as quadratic terms for precipitation, temperature, and wind, and a cubic temperature variable. We also interact temperature with wind and add hourly lags for temperature and precipitation. We include as further controls an indicator for day of week, an indicator for hour of day in order to control for special pollution patterns throughout the day (e.g. increased traffic during rush hours), and an indicator for the month of year for seasonal effects. We then take the residuals from this regression and aggregate them by year and district.

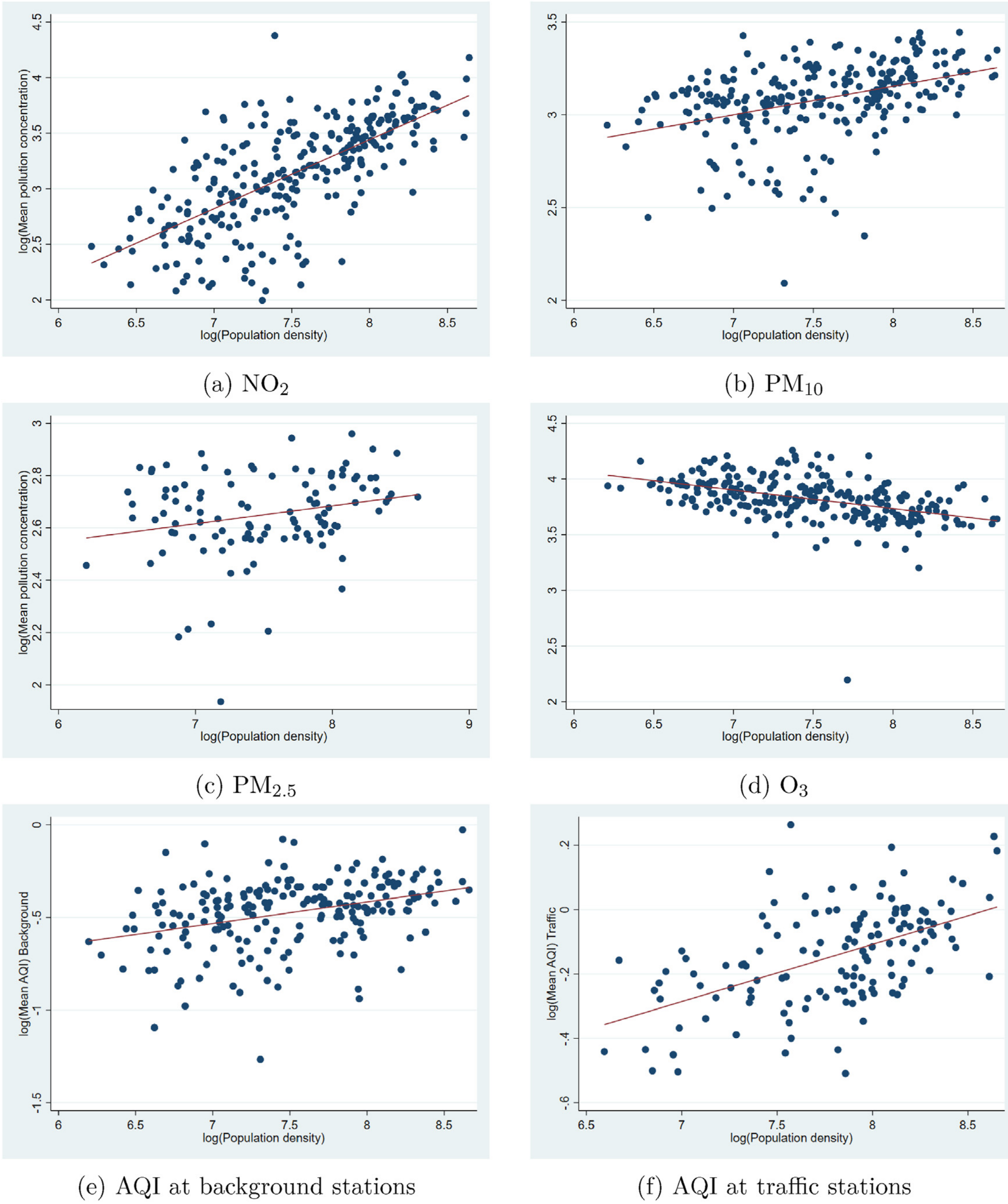
We then proceed with these residuals and start with a simple OLS model and regress pollution outcomes on a set of control variables. Our dependent variable, Y_{idt}, is the residual concentration level (for a particular pollutant) in year t at station i in district d.[29] Our main estimate of interest is population density, which is available in yearly intervals at the district level. Our first regression equation is

$$\ln(Y_{idt}) = \beta + \rho \ln(D_{dt}) + \gamma \mathbf{X}_{idt} + \alpha_t + \epsilon_{idt}. \quad (1)$$

Our main results stem from single measurement stations which are assigned to the closest weather station (see above). Emissions are regressed on a set of control variables **X**. Those are attributes of the monitoring station like the station area (urban, suburban or rural), the distance to the next large road, and station type (background, traffic, or industrial). Therefore, we explicitly control for the location of a measuring station and its type. We also control for the distance of an emission station to the center of the most densely populated municipality within a district. Effectively, our measure of pollution is then the pollutant

[28] It is not immediately clear whether this variable should be included in the regressions: on the one hand, the energy mix might itself be driven by population, so one might want to leave the presence of coal fired power plants out. On the other hand, part of the location of these plants may be driven by the exogenous presence of coal mines. We therefore include coal fired power plants only in the robustness section; as will be seen, including this variable does not affect our results.

[29] Alternatively, we could take the average of the monitor readings by district.



(a) NO_2 (b) PM_{10} (c) $PM_{2.5}$ (d) O_3 (e) AQI at background stations (f) AQI at traffic stations

Fig. 2. Scatter plots of log(pollution) with log(Population density) and linear fits.

concentration at the CBD. This should be a representative measure of city pollution. We also include year dummies α_t in order to control for business cycles and other time varying effects. District-level population density is denoted D_{dt}. Our parameter of interest is then ρ, which measures the elasticity of pollution concentration with respect to population density.

We can add economic controls at the district level like GDP, mean household income, and the unemployment rate. We can also control for whether the station lies within an environmental zone with a red, yellow or green badge. The share of green party voters is used as a control for the sorting of households with ‘green’ preferences into low- or high-emission cities. There is also the concern that the presence of coal-fired power plants might cause bad air quality in some regions. Therefore, we constructed an indicator which equals one if such a power plant is located within the same district as the monitoring station. Moreover, we calculated the distance of a measuring station to the closest coal-fired power plant.

In choosing whether to include control variables, we face two issues. On the one hand, leaving important drivers out of the regression will lead to omitted variable bias. On the other hand, some of these variables may be endogenous and therefore constitute “bad controls” that should be left out of the regression. For instance, income may be affected by density through agglomeration effects (even though the large empirical literature tends to find modest agglomeration economies, e.g. Combes et al., 2010). This also holds for most other potential controls. Green voting clearly may differ with a district’s urbanity and also responds to local pollution. Coal fired power plants may be present in densely populated districts with large energy demand. The presence of LEZs and the share of transit users are also likely affected by population density. Therefore, we choose to present our basic regressions with controls only for the urban/suburban/rural indicator, station type, distance to major street, and distance to the CBD. As a sensitivity check, we analyze in Sec. 5.2 how our results change when we successively add controls.

We cluster standard errors on the labor market region-year level in our OLS regressions. According to Cameron and Miller (2015, p. 333), the consensus is to be conservative and avoid bias by using "bigger and more aggregate clusters when possible, up to and including the point at which there is concern about having too few clusters". Compared to using clusters at district-year level, significance of the results does not change. However, we prefer using labor market regions as otherwise we have too few observations (monitoring stations) within some clusters.

OLS estimates would be unbiased and consistent if population density is not correlated with the error term, conditional on controls. However, this seems unlikely. For instance, densely populated cities may differ from less densely populated ones in their geography, industrial structure, or other unobserved variables that affect emissions.[30] Therefore, we also estimate long difference and fixed effects regressions of the form

$$\Delta \ln(Y_{idt}) = \rho \Delta \ln(D_{dt}) + \gamma \Delta \mathbf{X}_{idt} + \Delta \alpha_t + \Delta \widetilde{\epsilon}_{idt}, \tag{2}$$

where $\Delta \ln Y \equiv \ln Y_T - \ln Y_F$ and so on. We run a long-difference estimation where $t = F$ is 2002 and $t = T$ is the year 2015, while in other regressions, we include all years in the sample to estimate fixed effects. Our main long difference regressions control for unobserved heterogeneity at the district level, but we also consider long differences at the station level (see Appendix D). In addition to the controls described above, we again add year dummies α_t to the estimation. An alternative to long differences would be fixed effects regressions using the entire sample. We prefer the long-difference estimator since the yearly within variation of population density is small. However, we also perform fixed effects regressions and the results differ only slightly in the size of the estimated coefficients.

Long difference estimation will be unbiased if the unobserved heterogeneity that affects density and pollution is time invariant. However, if there are unobserved time varying factors which affect emissions and are correlated with density changes over time, the long difference estimates will be biased. For instance, it may be that household sorting leads to large cities getting 'greener' over time. In this case, density may still be correlated with the error term. Moreover, density and pollution may be simultaneously determined. For instance, households may migrate out of very polluted cities, which leads to endogeneity of population density. Moreover, as the within variation of density and pollution is relatively low, fixed effects estimates may suffer from imprecise estimates. Therefore, we also estimate instrumental variables (IV) regressions:

$$\ln(D_{dt}) = B_0 + B_1 \mathbf{X}_{idt} + B_2 Z_{dt} + \eta_{idt} \tag{3}$$

$$\ln(Y_{idt}) = \Gamma + \rho \ln(\widehat{D}_{dt}) + A_1 \mathbf{X}_{idt} + \widehat{\epsilon}_{idt} \tag{4}$$

Here, in the first stage regression (3), density is regressed on one or more instrumental variable(s) Z. The IV will be valid if the instrument strongly predicts density but is not correlated with the error term in the second-stage regression (4).

Like Combes et al. (2010), we use both historical population data and soil quality as instruments. The use of historical population data follows a long tradition starting with Ciccone and Hall (1996). We use the log of historical population density from1910[31] as instrument for current density.[32] Historical population data are relevant, since urban population tends to be strongly persistent over time. The exclusion restriction requires that historical density be correlated with current emission levels only through its effect on current density. We believe this to be the case, since pollution in the early 20th century was driven largely by industry. Today's urban pollution is much more driven by automobile traffic, which was close to non-existent in 1910. The German emperor Wilhelm II is purported to have said around 1900: "I believe in horses. Automobiles are just a phenomenon of temporary importance" (our translation). Furthermore, industry structures have changed dramatically over time. Therefore, it seems unlikely that historical population patterns should directly affect current pollution. We address this concern further below.

Following Combes et al. (2010), in addition to historical population densities, we instrument current population densities with data on soil characteristics. Some soil materials are better suited for construction to support a large number of households. Furthermore, in the past households were attracted to settle in areas with fertile land. Henderson et al. (2018) argue that agricultural variables are the most important drivers of agglomeration, especially in developed countries. Therefore, soil characteristics should be important determinants of historical and current population patterns.[33] For these variables, the exclusion restriction may be easier to justify (Combes et al., 2010). First, geology is largely determined by nature and should thus be independent of human economic activity. Second, since agriculture accounts for less than 5% of current employment, soil characteristics should not be important drivers of current pollution levels. We return to the issue of instrument exogeneity below and provide some additional tests in support of it.

Soil characteristics are taken from the European Soil Database (ESDB). For the choice of variables, we follow Combes et al. (2010) who look at French regions. We consider only variables that tend not to be influenced by human activity and therefore should be exogenous to it. In particular, we use soil characteristics that describe the mineralogy of the topsoil and the subsoil as well as the dominant parent material of the soil. The dominant parent material describes the bedrock of the soil, which is the underlying geological material. Mineralogy captures the presence of minerals in the different layers of soil. We also include information about the topsoil organic carbon content and the soil profile differentiation.[34] The last variable we use is the ruggedness of a district, which is the difference of the mean of maximum altitudes and the mean of minimum altitudes of all available values within a district.[35] More detail on the soil data can be found in Appendix C. In all of our instrumental variable regressions we cluster standard errors at the labor market region level. Since all our instruments are time invariant, we do not cluster at year level here.

[30] For instance, Stuttgart, one of the most densely populated cities, lies in a valley which makes it prone to high pollution concentrations.

[31] The authors would like to thank Uli Schubert from gemeindeverzeichnis.de for sharing his data on population in 1910.

[32] See, e.g. Koh et al. (2013) and Redding and Sturm (2008) who use similar historical data for Germany. Note that there is no consistent population data for earlier years covering *all* districts, so instead of using incomplete data going further back in time we choose 1910 to have a complete IV.

[33] Note, however, that soil characteristics are a narrower determinant of current population than historical population, see Combes et al. (2010).

[34] Due to the limited number of observations, we combine the high and medium categories of soil carbon contents into one category. The categorical values of topsoil mineralogy cannot be combined, as they are not ordinal. Thus, we replace the one value which occurs only once in the dataset with a missing.

[35] All variables that we consider as instruments have at least one category which is significant at the 10-percent level or higher in the first stage regression. Variables with no significant category (water capacity at the topsoil and the subsoil, depth to rock, soil erodibility class, and hydrological class) are not included in the regressions. However, including those variables does not alter the quality of the second-stage results, but leads to the instruments getting weaker and sometimes overidentified.

Table 2
OLS and IV regressions for NO_2 and PM_{10}.

	NO_2				PM_{10}			
	(1) OLS	(2) IV Density 1910	(3) IV Soil	(4) IV 1910 & Soil	(5) OLS	(6) IV Density 1910	(7) IV Soil	(8) IV 1910 & Soil
log(pop density)	0.280***	0.191***	0.292***	0.251***	0.0749***	0.104***	0.0662*	0.0787***
	(0.0141)	(0.0544)	(0.0600)	(0.0575)	(0.00691)	(0.0239)	(0.0366)	(0.0247)
Distance to CBD	−0.00321***	−0.00477***	−0.00299**	−0.00378**	0.000947***	0.00115	0.000809	0.000744
	(0.000447)	(0.00152)	(0.00139)	(0.00147)	(0.000297)	(0.000906)	(0.000937)	(0.000884)
Distance to Street	−0.105***	−0.102**	−0.103***	−0.101***	−0.0405***	−0.0350*	−0.0392**	−0.0333*
	(0.0115)	(0.0397)	(0.0377)	(0.0387)	(0.00648)	(0.0188)	(0.0199)	(0.0192)
Suburban	0.281***	0.281***	0.283***	0.279***	0.0761***	0.0702***	0.0782***	0.0728***
	(0.0169)	(0.0500)	(0.0501)	(0.0509)	(0.00948)	(0.0251)	(0.0263)	(0.0256)
Urban	0.445***	0.476***	0.443***	0.458***	0.140***	0.119***	0.144***	0.129***
	(0.0216)	(0.0678)	(0.0693)	(0.0716)	(0.0108)	(0.0324)	(0.0330)	(0.0315)
Industrial	0.0898***	0.0890**	0.0972***	0.0969***	0.136***	0.131***	0.139***	0.134***
	(0.0128)	(0.0368)	(0.0330)	(0.0345)	(0.0127)	(0.0364)	(0.0372)	(0.0374)
Traffic	0.648***	0.656***	0.647***	0.652***	0.255***	0.257***	0.257***	0.260***
	(0.0124)	(0.0402)	(0.0385)	(0.0402)	(0.00672)	(0.0186)	(0.0182)	(0.0187)
N	5575	5301	5547	5273	4648	4407	4620	4379
R^2	0.755	0.751	0.754	0.754	0.474	0.463	0.476	0.468
Districts	269	269	269	269	247	247	247	247
Soil Characteristics	No	No	Yes	Yes	No	No	Yes	Yes
First-stage F-statistic	–	318.5	11.79	46.72	–	383.3	10.93	48.64
Hansen p-stat	–	–	0.0681	0.0657	–	–	0.230	0.427

Standard errors in parantheses are clustered at labor market region - year (OLS) and labor market region (IV) level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.3.2. Threshold regressions

To test whether extreme values of PM_{10}, $PM_{2.5}$ or NO_2 correlate with population density, we run further regressions. We use the same basic approach as in Section 4.3.1, but now our dependent variable is a dummy variable which is one when the threshold value was violated and zero otherwise. Annual thresholds account for constant long-term exposure to air pollution. However, there might be elevated pollutant-concentrations throughout a year, which also tend to have more severe health effects. Thus, we furthermore examine whether densely populated areas tend to have more days with threshold violations (24-h means). Therefore, we created dummy variables which equal one when a station exceeded a predetermined number of days within a year. The outcome is therefore the probability of population density exceeding the pollution thresholds by a certain number of days within a year. The thresholds we look at are those set by the WHO air pollution guidelines (see Section 4.1.1). The number of days we choose are motivated by the number of days with threshold exceedances allowed by the EU. The hourly NO_2 threshold is allowed to be exceeded up to 18 times during a year and the PM_{10} threshold for 35 days. As in the EU (unlike the WHO) there is no short-term threshold for $PM_{2.5}$ there are also no allowed daily violations. Thus, we take the threshold value set by the WHO and as allowed number of days the same value as for PM_{10}. Furthermore, we look at the probability of violations on a certain number of days below the EU allowances (17, 14, and 9 days for NO_2, 34, 29, and 24 days for PM_{10} and for $PM_{2.5}$.) Local governments might try to take short-term measures to avoid illegal threshold violations, but still be subject to high pollution levels, so looking at threshold violations just below the allowances is of interest.

We use linear probability models (LPM) to estimate our outcomes of interest. With this approach, we can easily apply instrumental variable regressions. We think that the LPM does a decent job in estimating the probabilities, as the occurrence of transgressing the threshold is relatively dispersed over the sample. However, we also run probit regressions to account for potential non-linearities in the probability of transgressions (see Section 5.2).

5. Results

5.1. Basic results

OLS regressions. We present our basic cross-sectional OLS results in columns (1) and (5) of Table 2 and Table 3. The tables present coefficients for our parameter of interest, log of population density, as well as our basic controls (distance to CBD, distance to major street, whether the station lies in an urban or suburban area – rural is the reference category –, and the traffic and industrial station dummy – the reference category being background).

As shown by the OLS regression in column (1) of Table 2, the density elasticity of NO_2 concentration is 0.28 and the estimate is significant at 1%. The mean value of population density in 2015 was 2590.2 with a standard deviation of 1337.6. Thus, a one standard deviation increase in population density within a city increases the NO_2 concentration by 12.4 percent.

For PM_{10}, we find a smaller elasticity of 0.075, which is significant at 1% (column (5) of Table 2). A one standard deviation increase in population density increases the PM_{10} concentration by 3.2%.

For $PM_{2.5}$, the estimated elasticity is 0.035 which is significant at the 5% level (column (1) of Table 3). A one standard deviation increase in population density increases the $PM_{2.5}$ concentration by 1.48 percent. Note that the estimate is lower than the corresponding estimate for PM_{10}. However, the net of monitoring stations is both more recent and less dense, so the estimates for $PM_{2.5}$ are much less precise. Using a z-test, we cannot reject the hypothesis that the two coefficients are identical. Furthermore, below we reestimate the regressions using satellite data to measure pollution concentration. Since these cover the entire country, the estimates are much more precise. Interestingly, the IV estimate for $PM_{2.5}$ with satellite data is 0.08, just like the IV estimate for PM_{10} using station data.

Ozone concentration is negatively correlated with population density. This is probably due to the fact that the chemical prerequisites for

Table 3
OLS and IV regressions for $PM_{2.5}$ and O_3.

	$PM_{2.5}$				O_3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV Density 1910	IV Soil	IV 1910 & Soil	OLS	IV Density 1910	IV Soil	IV 1910 & Soil
log(pop density)	0.0353**	0.0820	0.0489	0.0204	−0.177***	−0.0931***	−0.217***	−0.143***
	(0.0161)	(0.0579)	(0.0658)	(0.0519)	(0.00945)	(0.0317)	(0.0503)	(0.0342)
Distance to CBD	0.00122*	0.00164	0.00137	0.000618	0.00230***	0.00354***	0.00172*	0.00279***
	(0.000651)	(0.00158)	(0.00166)	(0.00149)	(0.000315)	(0.00102)	(0.00102)	(0.000982)
Distance to Street	−0.0408**	−0.0431	−0.0341	−0.0274	0.0522***	0.0474**	0.0533**	0.0495**
	(0.0172)	(0.0394)	(0.0421)	(0.0440)	(0.00727)	(0.0241)	(0.0246)	(0.0243)
Suburban	0.124***	0.115*	0.135**	0.144**	−0.139***	−0.143***	−0.135***	−0.138***
	(0.0294)	(0.0617)	(0.0654)	(0.0648)	(0.0117)	(0.0347)	(0.0371)	(0.0349)
Urban	0.161***	0.123*	0.159*	0.180**	−0.193***	−0.219***	−0.180***	−0.204***
	(0.0330)	(0.0745)	(0.0820)	(0.0735)	(0.0138)	(0.0431)	(0.0484)	(0.0443)
Industrial	0.0619***	0.0659	0.0814**	0.0928**	−0.0754***	−0.0501	−0.0855***	−0.0651**
	(0.0210)	(0.0426)	(0.0403)	(0.0390)	(0.0116)	(0.0333)	(0.0286)	(0.0291)
Traffic	0.109***	0.109***	0.117***	0.112**	−0.231***	−0.238***	−0.232***	−0.238***
	(0.0173)	(0.0418)	(0.0423)	(0.0450)	(0.0164)	(0.0411)	(0.0380)	(0.0393)
N	795	758	780	743	3776	3588	3756	3568
R^2	0.254	0.227	0.256	0.239	0.445	0.426	0.440	0.440
Districts	109	109	109	109	251	251	251	251
Soil Characteristics	No	No	Yes	Yes	No	No	Yes	Yes
First-stage F-statistic	–	114.5	7.909	25.12	–	385.4	12.88	47.54
Hansen p-stat	–	–	0.0447	–	–	–	0.0124	0.00767

Standard errors in parantheses are clustered at labor market region - year (OLS) and labor market region (IV) level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4
OLS and IV regressions for air quality index.

	Background Stations				Traffic Stations			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV Density 1910	IV Soil	IV 1910 & Soil	OLS	IV Density 1910	IV Soil	IV 1910 & Soil
log(pop density)	0.135***	0.0694	0.224***	0.128***	0.135***	0.0912	0.145**	0.114*
	(0.0128)	(0.0436)	(0.0730)	(0.0399)	(0.0162)	(0.0840)	(0.0684)	(0.0596)
Unemployment share	−0.269	0.0167	−0.835	−0.388	0.394*	0.709	0.285	0.523
	(0.187)	(0.470)	(0.677)	(0.485)	(0.236)	(0.652)	(0.634)	(0.591)
Av. GDP	0.00360	0.0694	−0.0549	0.0265	−0.0195	0.0191	−0.0253	0.00596
	(0.0145)	(0.0484)	(0.0572)	(0.0466)	(0.0187)	(0.0651)	(0.0595)	(0.0576)
Av. Income	0.0896*	0.129	0.0131	0.0652	0.173***	0.145	0.149	0.110
	(0.0524)	(0.150)	(0.130)	(0.137)	(0.0630)	(0.166)	(0.152)	(0.162)
Green Voters	−0.281*	−0.0869	−0.745	−0.353	0.916***	1.091*	0.911	1.034*
	(0.169)	(0.434)	(0.608)	(0.444)	(0.180)	(0.593)	(0.554)	(0.539)
N	2142	2001	2137	1996	1147	1087	1139	1079
R^2	0.495	0.481	0.483	0.486	0.365	0.366	0.372	0.379
Depend. Var.	log(AQ)	log(AQ)	log(AQ)	log(AQ)	log(AQ)	log(AQ)	log(AQ)	log(AQ)
Districts	200	182	195	181	134	120	128	118
First-stage F-statistic		123.1	5.897	13.36		35.52	12.54	19.17

Standard errors in parantheses are clustered at labor market region - year level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

ozone formation are more favorable outside large cities.[36] The density elasticity of ozone concentration is −0.18. A one standard deviation increase in population density decreases the O_3 concentration by 7.1 percent. This result is interesting, as it shows that not all pollution is worse in denser cities.

So far, we have looked at the effect of population density on individual pollutants, but city residents will most likely be interested in the overall air quality, i.e. accounting for all pollutants at once. To assess how air quality is affected by density, we measure its effect on the AQI described in Section 4.1.1 above. Results are in Table 4. The results show a density elasticity of bad air quality of 0.14 for background stations and for traffic stations in the OLS regressions (see columns 1 and 5). A one standard deviation increase in population density increases the AQI by 5.8 percent. This is another important finding: we have seen above that NO_2 and PM concentrations increase, whereas ozone concentration decreases with density. However, using a common index, we find that overall air quality decreases with density.

IV regressions. We now turn to the IV regression results. To judge the relevance of the instruments, in Table A1 we regress population density on each of the instruments and report the R^2. Historical population density is the strongest predictor of current population density, while the strength of the explanatory power varies for the geological instruments. For example, soil differentiation and subsoil mineralogy by themselves explain only about 1–3% of the variation in current population density, while the carbon content in the soil and the dominant parent material explain between 15 and 28% of the variation. In our main regressions we will additionally present F-statistics and partial R-

[36] This is because nitrogen monoxide (NO), which is contained in car exhaust fumes, reacts with ozone to NO_2. Ozone is split into O_2 and NO_2 such that ozone pollution in city centers is significantly lower. On the other hand, the ozone precursors are transported out of cities by wind and contribute to the formation of ozone away from their actual sources. See https://www.umweltbundesamt.de/daten/luft/ozon-belastung# textpart-1.

squares in order to gauge the instruments' relevance.

In Table OA.1 in the Online Appendix, we report results from regressions of population density on our instruments individually and combined. Again, some of the geological instruments by themselves seem weak, with low values of R^2 and F-statistics. However, historical density as well as the geological instruments combined are strong instruments.[37] Therefore, our instruments are relevant in that they explain a large share of the variation in current population density.

IV results are shown in Tables 2 and 3. As argued above, we think of historical density as a stronger instrument, while perhaps the exclusion restriction is more easy to defend for soil characteristics. Below, for want of a simple decision criterion, we rely on the results using both sets of instruments; however, using only one of them does not change the interpretations dramatically in most cases.

According to the estimates, the density elasticity is 0.25 for NO_2, 0.08 for PM_{10}, and 0.02 for $PM_{2.5}$, although again, the latter is imprecisely estimated. For O_3, the IV estimate is −0.14. Finally, the IV estimate for the AQI is 0.13 for background and 0.11 for traffic stations. Thus, a one standard deviation increase of population density increases the NO_2 concentration by 11%, and the PM_{10} concentration by 3.3%. The O_3 concentration decreases by 5.8% for a one-standard deviation increase in density. Last, the AQI increases by 5.5% for background stations and 4.9% for traffic stations. In general, the IV results using soil characteristics as instruments are fairly similar to the OLS results, while there is a bit more variation if we use the historical density instrument. In summary, it seems that the bias from omitted variables in OLS regressions is small, a point also made by Combes et al. (2010).

The argument for the exogeneity of historical instruments is that agglomeration tends to persist, so large and dense cities of the past tend to be large and dense today. However, if enough time has elapsed, industry structures and other unobserved factors that might be correlated with current pollution levels should have changed sufficiently over time; therefore, the assumption that historical density does not affect current pollution other than through its effect on current density seems plausible. However, the concern remains that some unmeasured historical characteristic that is correlated with past density and persistent over time affects current pollution. For instance, cities that were large industrial centers and densely populated in the past may still contain a lot of dirty industry today.

Likewise, places with fertile soil have historically become dense settlements because they could sustain large populations. The exclusion restriction is that these characteristics do not affect current pollution directly, for instance because these places still contain a lot of agriculture that contributes to air pollution. Since agriculture makes up less than two percent of employment (see www.umweltbundesamt.de) and contributes around 10% to total air pollution, this seems plausible.[38]

To address the concern of potentially endogenous instruments, we include two additional tests. First, we control for historical shares of workers in industry and crafts and agriculture in our basic IV regressions. Results are in Table A2. We control for the share of workers in industry and crafts in the IV with historical density (odd columns), and for the share workers in agriculture in the IV with soil instruments (even columns). We find that the coefficients on population density change only slightly, especially for the historical instrument.

Second, we regress the current shares of employment in industry, agriculture, and manufacturing on past population density. The results are in Table OA.2. Interestingly, they show that cities with high historical density contain *less* industrial and agricultural employment today. Thus, it seems like over a century or so, structural change led large centers of industry to shift into services; likewise, historically dense cities today contain less agricultural employment. It seems, therefore, that this kind of structural change renders a correlation of historical population patterns with current pollution unlikely.

Moreover, the overidentification tests do not seem to indicate instrument endogeneity (see the Hansen p-statistics in Tables 2 and 3).

Fixed effects and long differences. Fixed effects regressions may be a proper response to time invariant unobserved heterogeneity that causes cities to be more or less dense and more or less polluted at the same time. For instance, if dense cities provide amenities which attract 'green' households and these households influence local pollution policies, the correlation of density and pollution might be driven by household selection. Using fixed effects at the district level could mitigate this selection bias. However, the within variation in density and pollution is much lower than the between variation, so fixed effects take out a lot of the interesting variation and the coefficient of interest is less precisely estimated. We present long-difference estimates for the years 2002–2015 as well as district fixed effects outcomes in Table 5. Fixed effects have the advantage of providing more observations (all years between 2002 and 2015), while the within variation of density and pollution is lower than for the long differences.

As Table 5 shows, the estimated coefficient on population density becomes insignificant in all but the NO_2 regressions. This seems to be because of the lower precision of the estimates due to the lower within variation of population densities. The coefficient in the NO_2 regression is 0.356 in the long difference regression, and 0.337 with district fixed effects, and both coefficients are significant at the 5 percent level.

We also ran station fixed effects regressions, which are presented in Table A3. The magnitude of the coefficient in the NO_2 regressions is very similar to the one in the OLS and IV regressions.

The NO_2 results point to the importance of car traffic for the production of air pollution in German cities. The recent discussion on threshold violations (see Section 5.3) and driving bans for Diesel cars underlines the current political dimension of this debate. In fact, when we rerun the long difference regressions by station type, we find a large and significant effect of density on NO_2 concentration for traffic stations only.[39] This supports the view that car traffic is the main driver of NO_2 concentration in dense cities.

5.2. Further results and robustness

In this subsection, we perform a number of robustness checks to see how sensitive the results are to various specifications and to shed light on some interesting issues. We first look at how inclusion of control variables affects the estimates. We then perform several variations on the definitions of cities and population density and use alternative pollution measurements using satellite data. Finally, we discuss several potential mechanisms that may drive our results.

Inclusion of controls. First, we check how sensitive the results are to the inclusion of controls. On the one hand, this may give some indication of whether our results are prone to suffer from omitted variable bias. On the other hand, we have several variables that may themselves be endogenous, but which may serve as mediating variables through which density affects pollution. This issue will be separately discussed below. We start with population density and year fixed effects as the only explanatory variables and successively add further control variables to the OLS regressions. Results are shown in Tables OA.3 to OA.6. Looking at the results for NO_2, we see that adding station-specific control variables (urban/suburban, distance to CBD, distance to major road, station type; column 2) cuts the coefficient on population density in half. Column (3) adds an indicator for the state (*Bundesland*) in order to control for state-specific policies. This reduces the coefficient size

[37] The results are shown only for the sample of NO_2 stations. We repeated these regressions for the subsamples of stations covering the other pollutants, but results do not differ significantly between them.

[38] More precisely, agriculture contributes 10% to NO_X emissions, 5% to $PM_{2.5}$ and 15% to PM_{10} emissions, see www.umweltbundesamt.de.

[39] Results are available on request.

Table 5
Long Difference estimations from 2002 to 2015 and Fixed effects estimations with all years.

	NO_2		PM_{10}		$PM_{2.5}$		O_3	
	(1) FE All years	(2) LD 2002-15	(3) FE All years	(4) LD 2002-15	(5) FE All years	(6) LD 2002-15	(7) FE All years	(8) LD 2002-15
log(pop density)	0.337**	0.356**	−0.0223	−0.101	0.308	0.615	0.254	0.358
	(0.133)	(0.178)	(0.0913)	(0.157)	(0.228)	(1.768)	(0.157)	(0.228)
Distance to CBD	−0.00995***	−0.00851***	−0.00359**	−0.00225	−0.000513	−0.00113	0.00500***	0.00254
	(0.00231)	(0.00268)	(0.00145)	(0.00233)	(0.00392)	(0.00590)	(0.00163)	(0.00232)
Suburban	0.345***	0.312***	0.0986***	0.0139	0.0311	0.0575	−0.153***	−0.141*
	(0.0589)	(0.0877)	(0.0364)	(0.0616)	(0.101)	(0.241)	(0.0576)	(0.0767)
Urban	0.574***	0.536***	0.188***	0.118*	0.127	0.168	−0.232***	−0.248***
	(0.0488)	(0.0527)	(0.0392)	(0.0620)	(0.102)	(0.222)	(0.0656)	(0.0798)
Industrial	0.131***	0.158***	0.127***	0.0474	0.0542	−0.0122	−0.0875*	−0.0876
	(0.0433)	(0.0411)	(0.0414)	(0.0513)	(0.0393)	(0.0810)	(0.0455)	(0.0821)
Traffic	0.717***	0.668***	0.276***	0.262***	0.263***	0.231**	−0.213***	−0.223**
	(0.0337)	(0.0404)	(0.0151)	(0.0238)	(0.0337)	(0.104)	(0.0468)	(0.0869)
N	5575	781	4648	545	795	135	3776	549
R^2	0.896	0.897	0.761	0.804	0.794	0.932	0.824	0.834
Districts	269	258	247	235	109	105	251	248

Standard errors in parantheses are clustered at labor market region level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

further.[40]

The presence of a coal-fired power plant may be a driving force for air pollution in some regions as pollution from these plants may be transported over long distances (Zhou et al., 2006). When we include an indicator for the existence of a coal-fired power plant in the district, the coefficient remains basically unchanged (column 4 of Table OA.3). In column (5) we replace the indicator variable with a measure of the distance to the closest coal-fired power station. This variable may be better able to capture possible spillovers from coal firing power plants. Monitors close to coal-fired power plants might be more affected than those farther away. The density coefficient is slightly reduced.

Adding control variables (log GDP per capita, log of average household income and share of unemployment in a district) in column (6) lowers the coefficient a little further. The density coefficient remains relatively stable in magnitude across the range of included control variables, and always remains highly significant. In summary, once we add a basic set of control variables which account for station-specific attributes, the coefficient does not change significantly anymore.

The picture is similar for the PM_{10} outcomes. Here, in particular adding state fixed effects reduces the density coefficient; the coefficient remains highly significant throughout all of the specifications.

For $PM_{2.5}$ (Table OA.5), the density coefficient becomes insignificant as soon as we add indicators for the presence of coal-fired power plants or when adding state fixed effects (columns 3 and 4).[41] However, comparing the sample distributions in Fig. 1 (Panel b) and in Fig. A2 (Panel a), we see that the $PM_{2.5}$ sample fails to cover many of the regions that are in the PM_{10} and the NO_2 sample. In particular, many of the densely populated areas like Berlin, the Stuttgart metropolitan area, and parts of the metropolitan areas of Hamburg and Munich are missing. It does seem, though, that the density effect on $PM_{2.5}$ is partly driven by the presence of coal-fired power plants in denser districts (Column 4 of Table OA.5).

For O_3, the picture is similar to the NO_2 outcomes (Table OA.6): as soon as we add basic controls, the coefficient is cut in half but remains highly significant throughout our specifications (except for the third column, where we add state fixed effects). In contrast to the other pollutants, however, the density coefficient is negative, so high density cities seem to suffer less from O_3 pollution.

Different definitions of cities and population density. Our next set of tests is designed to check whether our results are sensitive to the particular definition of population density. We first check for the effect of different definitions of what is the relevant spatial unit, before looking at different definitions of population density.

A common definition of a city is based on the economic relations between locations, usually measured by commuting flows. We therefore rerun our basic regressions for German labor market regions (LMRs, German *Arbeitsmarktregionen*) as defined by Kosfeld and Werner (2012). Similar to other concepts such as MSAs or Functional Urban Areas, LMRs are defined as collections of districts with significant commuting flows between them. There are 141 labor market regions, of which 128 contain at least one monitor. Results are shown in Table A4 (the IV regressions use historical density and soil as instruments). The results are very close to the estimates for districts. For $PM_{2.5}$, both the OLS and IV estimates turn significant.[42] Table OA.7 contains results using municipalities (*Gemeinden*) as spatial unit.[43] The general pattern that emerges is that across pollutants, both the OLS and IV coefficients are smaller in absolute size than those obtained with districts or LMRs.

In order to completely free ourselves from any unit definition, be it a city or some sort of administrative area, we used GIS software to create buffers of one and of 5 km around monitoring stations and captured population density within these areas. In Table OA.8 we see that this station specific population density measure leads to lower estimates the smaller the buffer size. For NO_2, PM_{10}, and O_3 regressions, the estimate doubles in size when increasing the buffer from 1 to 5 km.

While we cannot say for sure which spatial scale is most appropriate, we think that municipalities and similarly smaller sized buffers are probably not the 'correct' unit, since taking this approach neglects the economic density of nearby geographic units which affect pollution,

[40] One possible reason for the reduced effect may be that Berlin and Hamburg, Germany's two largest cities, are also states and the coefficient captures the within-state effect. Running regressions without these two states/cities leads to a coefficient of about 0.22.

[41] When we control for distance to the next postal code with a coal-fired plant, the coefficient remains marginally significant.

[42] Additionally, we ran regressions using the definition of functional urban areas (FUA) as described by Moreno-Monroy et al. (2020). Results obtained with FUAs are very similar to the ones using districts. Results are available upon request.

[43] In Germany, there are between 11.000 and 12.000 municipalities of which only a very small share contains a pollution measuring station (366 municipalities in our NO_2 sample and even less in the samples with the other pollutants). Furthermore, number and form of municipalities changed quite substantially over time. From 2000 until 2015, the total number was reduced from about 14.000 to about 11.000.

e.g. through commuting and other economic activities (such as power plants or industrial spillovers) that produce spillover pollution. In other words, the smaller the spatial unit, the larger will the disparity between the generation and the local exposure to pollution be. In our view, the interpretation of linking pollution exposure to *economic* density is thus probably best viewed at scales larger than the community level.

The next issue we look at is the definition of population density. So far, we have defined population density as total district population divided by total built up area. However, some papers have used other measures of agglomeration (see e.g. Ahlfeldt and Pietrostefani (2019) for a discussion). We therefore rerun our basic regressions with different density measures, see Table OA.9. In particular, instead of population density, we now use the population density over the entire area (instead of built up area only), total population or the total employment per km^2 (all in logs). As is to be expected, the results differ somewhat from our main results quantitatively but not qualitatively. Using the alternative population density measure (density over the entire district area) cuts the coefficients in half for all pollutants. The coefficient on population is a bit smaller than the one for density in the case of NO_2 but larger for PM_{10} and $PM_{2.5}$. Obviously, population can be large in a large district with low population density, so the interpretation of the coefficient here is somewhat different. Looking at employment density opens another angle on the pollution-density relation. While looking at population density emphasizes residential energy use and commuting at short distances, looking at employment density instead focuses on commuting over longer distances and possible agglomeration effects on industrial pollution. As seen in the third rows of each panel, the coefficients are again close to the ones from our baseline results, especially when looking at NO_2. The distinction between employment and residential density should be particularly important for smaller spatial units such as municipalities. Indeed, labor market regions are constructed so as to maximize commuting within the unit, so the distinction between residential and employment density becomes unimportant. Between models (OLS or IV with our different instruments) the coefficients are relatively stable for almost all of the independent variables we look at.

Another interesting question is whether the effect of density on pollution is driven by traffic or 'background' activities such as residential energy use or perhaps industrial fumes that disperse over the entire city area. In Table OA.10, we interact population density with the station type indicator. The density coefficient now corresponds to the effect of population density on pollution at background stations; it remains positive for NO_2 and particulates. Intuitively, we find that the density effect on air pollution seems more pronounced at traffic and industrial stations.[44]

For additional evidence, we test whether the density effect differs between growing and shrinking cities. Sluggish responsiveness of infrastructure and housing stock may imply that growing and shrinking cities have different density-pollution relationships. On the one hand, the gradient may be steeper in growing cities if road infrastructure does not keep up with traffic increase, which might lead to congestion and higher pollution. On the other hand, it might also be steeper in shrinking cities, if infrastructure and housing stock do not shrink in par with the population, which could imply higher energy use at given densities than in growing cities. Fig. OA.1 shows the coefficient of population density for cities where population increased between 2002 and 2015 and those were population decreased. With the exception of $PM_{2.5}$, we find that (in absolute terms), the effect of density on pollution seems to be smaller in growing than in shrinking districts.

Up to now, we have assumed the effect of density to be linear. This need not be the case, as an increase of density might affect pollution differently depending on the level of density. For instance, traffic congestion may only pick up if density rises so much that traffic volume outpaces infrastructure supply. Hence, we now test for potential non-linear effects of density in a stylized way. We divide the sample into population density quintiles and test whether the density effect is different between quintiles (assuming it is constant within quintiles). Results are shown in Fig. A3. For NO_2 and O_3, the results indicate that the density effect is driven by the denser cities: pollution increases for NO_2 and decreases for O_3 when moving up the density distribution. For PM_{10}, the effect of density increases when moving from the first to the second quintile, but stays constant thereafter; for $PM_{2.5}$, there is no clear cut relationship (as before, the estimates are more noisy than for the other pollutants).

As a final robustness check, we rerun our regressions using satellite data instead of official monitor readings as our measure of pollution. Satellite readings are available for NO_2 and $PM_{2.5}$. These data have been used in many recent studies (e.g., Freeman et al., 2019; Achakulwisut et al., 2019). Compared to official monitor readings, the satellite data contain pros and cons. On the one hand, they are surely subject to measurement error, as pollution is not directly measured but inferred indirectly from related physical observation (e.g. aerosol optical thickness, which is inferred from the way the atmosphere reflects and absorbs visible and infrared light). On the other hand, the satellite data are available for grids with resolutions of 0.1° by 0.1° (approximately 10 km × 10 km at the equator) for NO_2 and 0.01° by 0.01° (1 km × 1 km) for $PM_{2.5}$, instead of the monitors which in our case are only placed in certain cities, but don't cover the entire country. Related to this, monitors may be placed endogenously in high pollution/high density locations. While these concerns should be mitigated by our long difference and IV estimates, it is still interesting to check how monitor and satellite based results match up. The correlation between the pollution measurements with station readings and satellite data is 0.4 for NO_2 and 0.7 for $PM_{2.5}$. More detail on the satellite data is found in Appendix C.

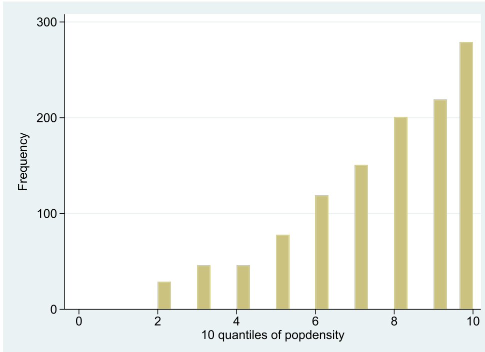
Table A5 shows the results. For $PM_{2.5}$, the satellite based results are close to our main results based on monitor readings. They are also more precisely estimated, which shows the upside of the much more complete coverage of satellite data. For NO_2, the results are similar to the baseline, but show a somewhat larger effect of density on emissions.[45]

We note, however, that satellite and monitor data are not directly comparable because of the differences in sample coverage. We therefore restrict the satellite sample to the sample of monitoring stations. That is, we now keep only grid cells (or districts) which have at least one monitoring station. We only show the OLS estimates without controls in Table OA.11. The table shows that the density coefficient is significantly higher with the station readings than with the satellite data. For $PM_{2.5}$, there is a smaller difference between the satellite and station estimates, which is not, however, significant. Since, if available, station data are more accurately measured, we prefer the estimates using station data for NO_2. For $PM_{2.5}$, since results are similar between station and satellite data, but the station net is not well developed, satellite data are probably preferable. Overall, however, the results using satellite data are broadly in line with our baseline results.

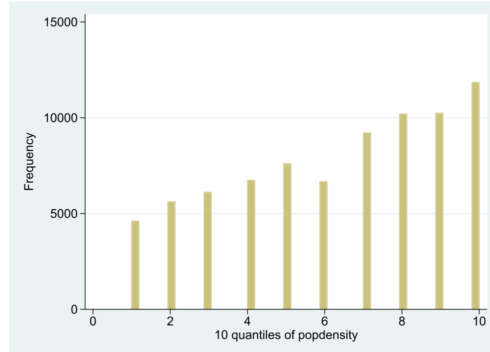
Discussion of possible mechanisms. We now discuss some potential mechanisms that may be responsible for our findings. Urban economic models like Borck and Pflüger (2019) analyze how urban pollution is driven by industrial and agricultural production, transport, and residential energy use. An alternative mechanism may be the sorting of "green" residents into cities. We therefore want to discuss what we can learn about potential channels from our analysis. One by one, we include the following variables. First, we look at the share of public transport users and car density. Second, we consider industrial composition. Third, we add some variables meant to measure the 'greenness'

[44] Results for O_3 show that the effect is not significantly different at traffic stations compared to background stations but is significantly lower at industrial sites. Results are available upon request.

[45] Carozzi and Roth (2019) also compare the two methods for $PM_{2.5}$ and find lower estimates of the density elasticity with monitor readings.



(a) Transgressions of annual NO_2



(b) Annual transgressions of daily PM_{10}

Fig. 3. Histograms of threshold transgressions by deciles of population density.

of cities: green space area, an indicator for environmental zones and the share of green party voters. If the density coefficient changes significantly, the variable may be viewed as a driver of the density-pollution gradient. Results are in Tables OA.12–OA.15. Panel A contains the regression results with and Panel B those without the corresponding variable.[46]

In the first two columns of Table OA.12–OA.15, we add the number of public transit users (col. (1)) and number of cars (col. (2)) in a district as controls. Denser districts have both more transit users and more cars (although the number of cars per capita is lower). Intuitively, we find that NO_2 and PM_{10} pollution is positively correlated with the number of cars and negatively with the number of transit users. Hence, commuting by car contributes to the positive density effect, while commuting by transit reduces it.[47] For O_3, the picture is again the opposite to NO_2: O_3 is reduced by the number of cars and increased by the number of transit users.

We now look at the number of buildings as a proxy for residential energy use. The results are in the last column of Table OA.12–OA.15. Denser districts obviously have more buildings. As the tables show, buildings are positively correlated with NO_2 and PM_{10} emissions. Hence, more residential energy use in denser cities seems to contribute to higher pollution.[48] Again, the opposite is true for O_3.

Next, we look at industry composition. It might be that dense cities are more polluted because they are specialized in dirty industries. To investigate this mechanism, we add the employment shares of industry types (industrial production, construction, agriculture, financial industries, public sector, and trade) to our regressions. The sectors that are most polluting are industry, construction and agriculture, whereas services and public sector should be much less polluting. We add the share of 'dirty' and 'clean' industries in total production separately as one block each (see cols (3) for 'dirty' and (4) for 'clean' industries). We can distill the following: at least for NO_2 and PM_{10}, controlling for 'dirty' industries increases the density coefficient, while employment in these industries increases pollution. In fact, denser cities are less specialized in dirty industries and more in relatively cleaner ones. Hence, industry composition would actually seem to lead to pollution falling with density. For the 'clean' industry shares, we find small effects on the density coefficients, so they don't seem to strongly affect the pollution-density relationship.

Next, we add some variables measuring the 'greenness' of cities in a physical sense as well as in the sense of the composition of city population. Our first variable is the total area of green space in a city. If densely populated cities have less green space, pollution concentration may be higher than in less densely populated cities, as trees and plants may capture or filter air pollution and lower temperatures. On the other hand, plants and trees emit VOCs, which are precursors to O_3. The results are shown in column (5) in the respective tables. Green space increases O_3 and decreases NO_2 concentration; this is consistent with the fact that plants emit VOCs which react with NO_2 to form O_3. Since, in fact, denser districts have more green space, controlling for green space slightly increases the density coefficient for NO_2, while decreasing the density coefficient for O_3. The effect of green space on PM_{10} is rather small.

Second, we add an indicator variable for whether or not a station lies in an LEZ. Cities may designate LEZs within their jurisdiction that only cars with specific colored badges may enter. The designation is a political decision and hence may mirror the environmental preferences of the residents. In column (6) of Table OA.12–OA.15, we see that LEZs do not alter the density coefficient even though they are mostly positively correlated with pollution concentration.[49]

Lastly, we consider whether the density effect might be mediated by sorting of households according to environmental preferences. For instance, families might move to less dense and greener locations to avoid adverse health effects for their children. Conversely, cities may attract "green" voters with a strong preference for the environment. We follow the last track and include the share of green votes in our regressions (col 7). For particulates and O_3, we find that the share of green voters reduces pollution. Since the green voter share is positively correlated with density, this reduces the density coefficient. This would imply that selection of green voters into cities makes denser cities greener.[50]

In summary, we find that the density effect is not driven by industry composition or the composition of the population in denser cities. Dense cities seem to have cleaner industries and are inhabited by 'greener' residents. However, consistent with our simple model, they have more

[46] This is useful to compare the sample composition when adding the respective control of interest. We also redid the exercise with IV estimates, and the results are very similar to the OLS ones. These results are available upon request.

[47] To get a complete picture, we would have to know how total vehicle kilometers for driving and transit change with density, as well as the emissions intensity of the two modes. This is beyond the scope of the paper, but we conjecture that while denser cities have a higher transit share and lower car density, they still have more traffic in total, which contributes to pollution even though the average commute may be cleaner than in less densely populated cities.

[48] Inuitively, building density is also larger in denser cities. Since this implies a higher energy efficiency Borck and Brueckner (2018), dense buildings flatten the density-pollution gradient.

[49] Of course, this may be due to reverse causality: if pollution is high, political pressure for introducing an LEZ mounts.

[50] Interestingly, for NO_2, the green vote share seems to positively correlate with pollution. However, once we control for the vote shares of other major parties (CDU, SPD, the Left and FDP), the outcome mirrors that for the other pollutants.

Table 6
Probability of transgression of annual thresholds.

	NO2		PM10		PM2.5	
	(1) LPM	(2) LPM IV	(3) LPM	(4) LPM IV	(5) LPM	(6) LPM IV
log(pop density)	0.148***	0.143***	0.100***	0.0866***	−0.0128	−0.0229
	(0.0232)	(0.0313)	(0.0214)	(0.0264)	(0.0153)	(0.0164)
distance to CBD	0.00143*	0.00136	0.00138	0.000820	0.000380	0.0000524
	(0.000818)	(0.000841)	(0.000993)	(0.00102)	(0.000809)	(0.000864)
suburban	−0.0376*	−0.0311	0.218***	0.207***	0.185***	0.136***
	(0.0191)	(0.0199)	(0.0403)	(0.0414)	(0.0643)	(0.0497)
urban	−0.0378	−0.0303	0.275***	0.264***	0.206***	0.158***
	(0.0244)	(0.0285)	(0.0459)	(0.0474)	(0.0710)	(0.0558)
industrial	−0.0000506	−0.00415	0.260***	0.255***	0.0395	0.0403*
	(0.0206)	(0.0208)	(0.0411)	(0.0426)	(0.0272)	(0.0225)
traffic	0.549***	0.563***	0.311***	0.316***	0.0124	0.0142
	(0.0427)	(0.0457)	(0.0303)	(0.0320)	(0.0140)	(0.0120)
N	5663	5341	4812	4520	795	743
R^2	0.494	0.507	0.421	0.421	0.240	0.179
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Districts	266	245	244	225	108	101

Standard errors in parantheses are clustered at labor market region level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7
Probability of transgressing thresholds by specific number of days.

	NO2			PM10			PM2.5		
	(1) >17	(2) >14	(3) >9	(4) >34	(5) >29	(6) >24	(7) >34	(8) >29	(9) >24
log(pop density)	0.0255*	0.0277**	0.0312**	0.0503***	0.0625***	0.0546**	−0.0334	−0.00681	−0.00931
	(0.0132)	(0.0138)	(0.0147)	(0.0171)	(0.0196)	(0.0224)	(0.0389)	(0.0348)	(0.0276)
distance to CBD	0.000295	0.000330	0.000354	0.000840	0.000633	0.000476	0.000367	0.000719	0.00124
	(0.000242)	(0.000247)	(0.000282)	(0.000701)	(0.000733)	(0.000868)	(0.00219)	(0.00148)	(0.00135)
suburban	−0.00563	−0.00595	−0.00730	0.0135	0.0251	0.0519*	0.376***	0.316***	0.275***
	(0.00526)	(0.00550)	(0.00575)	(0.0200)	(0.0240)	(0.0268)	(0.0844)	(0.0807)	(0.0787)
urban	−0.00938	−0.0100	−0.0117*	0.0148	0.0246	0.0579*	0.394***	0.341***	0.311***
	(0.00591)	(0.00620)	(0.00640)	(0.0225)	(0.0285)	(0.0314)	(0.0872)	(0.0853)	(0.0768)
industrial	0.00835	0.00858	0.00789	0.114**	0.129**	0.167***	0.0623	0.113**	0.0559
	(0.00627)	(0.00656)	(0.00705)	(0.0452)	(0.0542)	(0.0637)	(0.0543)	(0.0436)	(0.0358)
traffic	0.0439**	0.0470**	0.0592***	0.239***	0.270***	0.306***	0.161***	0.140***	0.0699**
	(0.0176)	(0.0182)	(0.0198)	(0.0219)	(0.0245)	(0.0236)	(0.0465)	(0.0404)	(0.0317)
N	5663	5663	5663	4817	4817	4817	795	795	795
R^2	0.053	0.057	0.063	0.268	0.298	0.333	0.311	0.253	0.219
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Districts	269	269	269	247	247	247	109	109	109

Standard errors in parantheses are clustered at labor market region level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

total commuting and more total residential energy use. Our interpretation of the results is that denser cities are thus more polluted, even though each resident may produce lower emissions due to higher transit shares and more efficient energy use.

5.3. Threshold results

We now turn to the analysis of threshold violations. These have been the primary focus of recent policy debates, as cities and national governments in Germany and other European countries have been sued for violations of legally binding thresholds.

For a first visual impression, Fig. 3 shows the number of transgressions of the daily mean threshold of PM_{10} (50 μg/m^3), and the NO_2 annual mean (20 μg/m^3)) by population density decile. The histograms suggest a clear positive association between density and threshold transgressions.

In Table 6, we present results for the probability that the yearly mean was exceeded for NO_2, PM_{10} and $PM_{2.5}$. We concentrate here on linear probability models (LPM), again using the historical and the soil IVs in some specifications. For $PM_{2.5}$, there is no significant relation between density and annual threshold violations. For NO_2 and PM_{10}, in contrast, all results are positive and highly significant. The probability that the annual NO_2 threshold of 40 μg/m^3 is transgressed is significantly higher in more densely populated areas. Coefficients (except for the one when using soil characteristics as instruments) are similar in NO_2 and PM_{10} regressions. We also repeat these estimations using probit IV models and get very similar results (see Table OA.16).

Results for the transgressions of the 24-h mean are shown in Table 7. The table shows that the probability of exceeding the 24-h threshold more than 17 days for NO_2 is significantly higher in denser areas, even though the point estimate is relatively small, at 0.026. The lower we set the number of days, the higher the coefficient. In the case of PM_{10}, the probability of exceeding the threshold for 17 days is also significantly higher in denser cities with a point estimate of 0.05. For $PM_{2.5}$, we find an insignificant effect of density on threshold violations (note again, however, the smaller sample size). As shown in Table OA.17, using a probit model does not change the results.

In summary, the evidence suggests that threshold violations occur more frequently in more densely populated cities.

6. Conclusion

In this paper, we have used panel data for German districts to estimate the effect of population density on air pollution. Our theoretical model predicts that denser cities should have higher pollution concentration, although there are some countervailing forces. The evidence to date has been largely inconclusive. To mitigate concerns about unobserved heterogeneity and omitted variables, we have used both long difference regressions and instrumental variables. Our preferred estimates come from the IV regressions, where we instrument population density with historical population and/or soil characteristics. We find that increasing population density by one percent increases NO_2 by 0.25 percent and PM_{10} by 0.08 percent. The results for $PM_{2.5}$ are less precisely estimated but of similar magnitude than those for PM_{10}. For O_3, we find denser cities have lower concentrations, with an elasticity of −0.14. Air quality as measured by the aggregate AQI decreases with population density, with an elasticity of 0.14.

The study thus contributes to our knowledge about the economic costs of agglomeration. The benefits of agglomeration due to labor market pooling, spillovers, matching etc. are by now well documented. However, there is much less robust evidence on the costs of agglomeration.[51] Thus, our study makes some headway towards a more complete picture of agglomeration benefits and costs. This seems important for urban policies. In Appendix B, we use a simple numerical example which shows that, based on our estimates, local pollution may reduce optimal city size by 7%. Knowledge of the elasticity of pollution with respect to population leads to a more complete understanding of the benefits and costs of agglomeration.

As far as we know, together with Carozzi and Roth (2019), this is the only study that seriously tries to estimate the causal effect of population density on pollution. More evidence from other countries surely will add to a more complete picture about this issue. For instance, whether or not population and pollution interact differently in developing and developed countries seems like an interesting and important question.[52] More research on the interaction of urban structure and pollution thus seems warranted.

CRediT authorship contribution statement

Rainald Borck: Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Philipp Schrauth:** Data curation, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no conflict of interest concerning this paper.

Appendix

A. A model

Consider a circular, open monocentric city where N residents commute to the CBD for work. A household living at x km from the CBD incurs round-trip commuting costs tx. Household utility is $v(c,q)=c^{1-\alpha}q^{\alpha}\mathcal{P}^{-\beta}$, where c is non-housing consumption, q consumption of housing floor space in square meters, and $\mathcal{P}$ is the concentration of local pollution in the city. Households are completely mobile in the city, so they achieve utility level u regardless of their location.

The household maximizes utility subject to the budget constraint, $w = c - tx + pq$, where w is wage income and p the price of housing per sq meter. Maximizing utility subject to the budget constraint gives the household's optimal housing demand $q=\alpha u^{\frac{1}{\alpha}}\mathcal{P}^{\frac{\beta}{\alpha}}(y-tx)^{1-\frac{1}{\alpha}}$, and the bid rent, i.e. the maximum willingness to pay per unit of housing floor space, $p=u^{-1/\alpha}(y-tx)^{\frac{1}{\alpha}}\mathcal{P}^{-\frac{\beta}{\alpha}}$.

Housing floor space is produced by profit maximizing developers, using capital K and land L as inputs. We assume a Cobb-Douglas production function written in intensive form $h = S^{\theta}, \theta > 0$, where $S \equiv K/L$ is structural density (capital deployed per unit of land) and h is the amount of floor space per unit of land. We normalize the price of capital to one. The developer maximizes profits per unit of land

$$\pi = S^{\theta} - S - R,$$

where R is the land rent paid to (absentee) landowners. Solving the developers' problem gives structural density, $S=\theta^{\frac{1}{1-\theta}}u^{\frac{1}{\alpha(\theta-1)}}(y-tx)^{\frac{1}{\alpha-\alpha\theta}}\mathcal{P}^{\frac{\beta}{\alpha(\theta-1)}}$, and the land rent function at distance x, $R=\left(\theta^{\frac{\theta}{1-\theta}}-\theta^{\frac{1}{1-\theta}}\right)u^{\frac{1}{\alpha(\theta-1)}}(y-tx)^{\frac{1}{\alpha-\alpha\theta}}\mathcal{P}^{\frac{\beta}{\alpha(\theta-1)}}$.

We consider a small open city. Residents are freely mobile between the city and the rest of the economy. Letting u be the exogenous utility level that can be attained in the rest of the economy, the equilibrium is defined by the two equations

$$R(\overline{x},u,\mathcal{P}) = R_A \tag{A.1}$$

$$\int_0^{\overline{x}} \gamma D(x,u,\mathcal{P})2\pi x dx = N, \tag{A.2}$$

where $\overline{x}$ is the distance from the city border to the CBD and R_A is the agricultural land rent and γ denotes the share of developable land at any distance x. $D=\frac{h(x,u,\mathcal{P})}{q(x,u,\mathcal{P})}$ is the population density at distance x from the CBD.

Solving (A.1) and (A.2) gives the endogenous city border $\overline{x}$ and number of residents N.[53]

Pollution is composed of emissions from commuting, $\mathcal{C}$, and residential energy use, H, weighted by the respective emissions factors. Transport emissions are assumed to be proportional to the aggregate commuting distance, and residential emissions are proportional to total housing floor

[51] See Combes et al. (2018) for a recent study on the costs of agglomeration implied by high land and housing prices. The interpretation of these costs is different however, as long as land and housing markets are competitive.

[52] See Borck and Schrauth (2019).

[53] We use the following parameter values: $\gamma = 0.75$, $e_H = e_C = 1$, $r_A = 50{,}000$, $w = 50{,}000$, $t = 500$, $\beta = 0.05$, $\theta = 0.75$, $\alpha = 0.25$. These values are similar to those used in Borck and Brueckner (2018).

space in the city. Letting the emissions intensities of commuting and housing be e_C and e_H, total emissions are

$$E = e_C\mathcal{C} + e_H H \tag{A.3}$$

$$\mathcal{C} = \int_0^{\overline{x}} x\gamma D(x)2\pi x dx \tag{A.4}$$

$$H = \int_0^{\overline{x}} \gamma h(x)2\pi x dx. \tag{A.5}$$

Finally, assume for simplicity that the concentration of air pollution is given by total emissions divided by land area.[54] Then concentration is given by $\mathcal{P} = E/(\pi\overline{x}^2)$.

How then does pollution concentration change with population density? To answer this question, we vary the parameter γ. For instance, government may increase γ by more liberal zoning policies (e.g. increasing the floor-area ratio by allowing more housing to be built per sq meter of land). When γ increases, the city shrinks spatially as the city border $\overline{x}$ moves inward, for given population. For given population, this has two effects on residents' utility: First, utility would increase, since housing has become less scarce. Second, however, pollution concentration rises: first, reduced competition for land raises aggregate housing consumption and residential energy use, and second, since the city shrinks, average and total commuting distance falls. The combined effect – assuming equal emissions coefficients on residential and transport emissions – is an increase in emissions, and an increase in concentration, since total emissions now diffuse over a smaller area.[55]

In our simulation, we find that residents' utility increases, as the housing effect dominates the increase in pollution concentration.[56] This will lead to in-migration from the rest of the economy, which increases the number of residents and the urban boundary $\overline{x}$, although not to its previous level. As a result of the increased population level, the effect on pollution concentration is reinforced, and density rises as well. Our simulation shows that the end result in the open city is that the increase in γ increases density and pollution concentration. The positive relation between concentration and density emerging from the model is shown by the upper blue curve in Fig. A1.

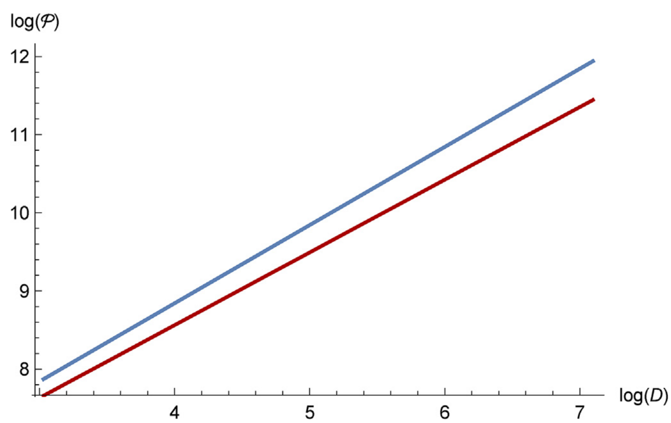


Fig. A.1 Population density and pollutant concentration

Extension. We now examine a couple of extensions that affect the relation between density and pollution. First, denser cities typically have higher mode shares for public transport because of economies of scale and traffic density. Second, denser cities have taller buildings that are more energy efficient. We include these two aspects in a simple reduced form fashion. In particular, we assume that transport emissions fall with average city density because density shifts transport mode choice towards cleaner public transit. Second, we assume that because of higher energy efficiency, density reduces the emissions associated with residential energy use.[57]

We amend eqs. (A.4) and (A.5) in the following way:

$$\mathcal{C} = \int_0^{\overline{x}} x\gamma D(x)dx\overline{D}^{-\kappa} \tag{A.6}$$

$$H = \int_0^{\overline{x}} \gamma h(x)dx\overline{D}^{-\mu}, \quad \kappa, \mu > 0, \tag{A.7}$$

[54] In reality, concentration is given by emissions per cubic meter of air, but we can slightly simplify by assuming all pollution is at ground level and thus concentration equals emissions over land area.
[55] If we increase the emissions coefficient for transport emissions sufficiently, we find that raising γ decreases aggregate emissions, but concentration still increases.
[56] This will be true as long as β, the parameter which governs the strength of pollution damage, is not too large. We choose a value similar to recent literature here, but even a 3-fold increase would not change our results.
[57] Borck (2019) contains a model with pollution and mode choice without scale economies. Borck and Brueckner (2018) use a microfounded model where energy use is related to a building's surface, which implies that more densely populated locations with tall buildings are more energy efficient.

where $\overline{D} = N/(\pi\overline{x}^2)$ is the (endogenous) average city density, and κ and μ are the density elasticities of transport emissions and residential energy emissions.

Ahlfeldt and Pietrostefani (2019) find that both transit mode share and energy efficiency rise with density, with both elasticities being about −0.07. Setting $\kappa = \mu = 0.07$ produces the red curve in Fig. A1. Thus, the relation between density and pollution remains positive but is flattened compared to the benchmark case where density economies are absent. In fact, the relation will remain positive as long as $\kappa, \mu < 1$. Given the estimates in Ahlfeldt and Pietrostefani (2019), this restriction seems very likely to hold.

B. Optimal city size

Consider a simplified version of the model above, where we abstract from housing construction. We modify the model in two respects: first, we assume that there is congestion, which implies that commuting costs per km are given by $t = N^{\psi}$. Second, we follow the literature on agglomeration economies and assume that the wage is given by N^{δ}. Suppose that pollution concentration can be described as in our empirical model by the relation $C = N^{\rho}$. Solving the model (using $\alpha = 1/4$) gives the following functional form for utility as a function of population size:

$$\log u = (\delta - \beta\rho)\log N - \frac{1}{4}\log\left(\frac{256}{27}\left(N^{1+\psi} + R_A\right)\right). \tag{A.8}$$

Suppose that $\beta = 0$, so households don't care about pollution. Furthermore, in line with the large empirical literature on agglomeration, let $\delta = 0.05$ (Combes and Gobillon, 2015), and, following Duranton and Puga (2019), let $\psi = 0.07$. Then, maximizing u with respect to city size gives $N^* = 6235$.

Suppose, however, that $\beta = 0.02$, and let $\rho = 0.15$ (between our estimates for PM_{10} and NO_2, and close to our coefficient for air quality).[58] Then, we find an optimal city size of $N^{**} = 5810$, 93% of N^*. Hence, pollution can significantly affect the balance of agglomeration benefits and costs.

C. Data

Weather. Since weather and emission stations are usually not at the exact same spot, we have to match emissions and weather stations such that we get the most accurate information about the weather at each emission station. Following the approach of Auffhammer and Kellogg (2011), for each emission station we searched for the ten closest weather and precipitation stations within a range of 50 km and a maximum station altitude difference of 200 m.[59] Out of those stations, we choose a primary station which is the closest weather or precipitation station to the emission station with at least 50 percent of hourly observations non-missing. All emission stations that could not be assigned a primary station were deleted from the sample. Throughout a year, there are gaps between recordings such that many weather and precipitation stations do not have a full record of observations. Such missing observations were imputed by regressing the non-missing values of, say, sunshine on the sunshine records of all the other adjacent stations. The estimated coefficients of those other stations were then used to impute values for missing observations.

About 80 percent of particulate matter and nitrogen dioxide emission stations were matched to the closest available weather station and less than four percent (PM_{10}) and two percent (NO_2) of emission stations were matched to a weather station ranked 5th or higher regarding the ranking of distance between the two station types. In both cases (PM_{10} and NO_2), less than 1 percent of emission stations could not be assigned a weather station.

Historical industry data. To construct the historical data for workers in industry and crafts, we proceeded as follows. We had maps for administrative units now and in 1925 and for 1925 the total number of workers in industry and crafts as well as the total population of a historical district. Due to the fact that administrative assignment changed over time, we had to assign historical administrative units to current units. If the historical area matched with current districts by more than 60 percent of the area, those areas were assigned the recent district. In many cases this is true for more than one historical district. For example, southern and northern Dithmarschen correspond to the current Dithmarschen. In these cases, we just summed the number of workers and the number of inhabitants in 1925 and assigned the sum to the current administrative unit. From these variables we then calculated the shares of workers in industry and crafts over the whole resident population. A number of current districts could not be assigned to workers because there were no historical districts matching by at least 60 percent of the area. This is true for example for Wolfsburg, a city that was established after 1925 and did not exist back then. Other cases like Mainz or Worms were larger districts in the past and were assigned as district-free cities after 1925. In such cases, the recent district almost completely lies within a historical district and we assigned the value of the respective historical district. As these are only relatively few cities and districts, we performed this matching by eyeballing the maps and looking which area fits best to the current district.

Geology. We use the same 12 variables from the European Soil Database (ESDB) used by Combes et al. (2010).[60] The data comes in raster format of 1 km × 1 km rasters, which we aggregate to the district level. For each district we use for instance the value of the dominant parent material which occurs most often within the district. Especially in urban areas like Berlin, we need to impute some of the values because of the lack of information in the data. In such cases, the dominant value often is described as a non-soil or just missing. In these cases we use the second most common value occurring within the district. The variables we use describe the mineralogy of the topsoil and the subsoil as well as the dominant parent material of the soil at different levels of aggregation. The dominant parent material describes the bedrock of the soil, which is the underlying geological material. At the broader level of aggregation, these are e.g. sedimentary rocks, igneous or metamorphic rocks, while the finer level of aggregation further classifies them. For instance, sedimentary rocks may consist of different types of limestone (hard, soft, marly, chalky etc.), marlstone or other types of stones. Mineralogy captures the presence of minerals in the different layers of soil (the topsoil being usually 5–15 cm deep and the subsoil being the intermediate layer between the topsoil and the bedrock).

We also include information about the water capacity of the topsoil (from low to very high) and the subsoil (from very low to very high), the depth to rock (from shallow to very deep), the soil erodibility class (from very weak to very strong), the topsoil organic carbon content (from low

[58] The value of 0.02 is close to the value used by Borck and Tabuchi (2018) based on a calibration to the social cost of carbon.

[59] There are many more precipitation stations in Germany (more than 4000) than stations which provide information on all other weather variables other than rainfall and snowfall (a little more than 700). This is why we separately merged precipitation and weather stations to each emission station.

[60] These data can be freely downloaded for research purposes from the European Soil Data Centre (Panagos et al., 2012).

to very high), the soil profile differentiation (no differentiation, low and high differentiation) and the hydrological class, which consists of four categories describing the circulation and retention of underground water. The last variable we use is the ruggedness of a district, which is calculated as the difference between the mean of maximum altitudes of all the rasters within a district and the mean of minimum altitudes across all rasters within the same district.

We include the information on mineralogy, hydrological class and parent material as dummies in the regressions. All other variables, which differ in the quality of a characteristic (e.g. from low to high) remain in their continuous form. All variables are included as dummies in the regressions, except for ruggedness, which is the only continuous variable among the soil characteristics.[61]

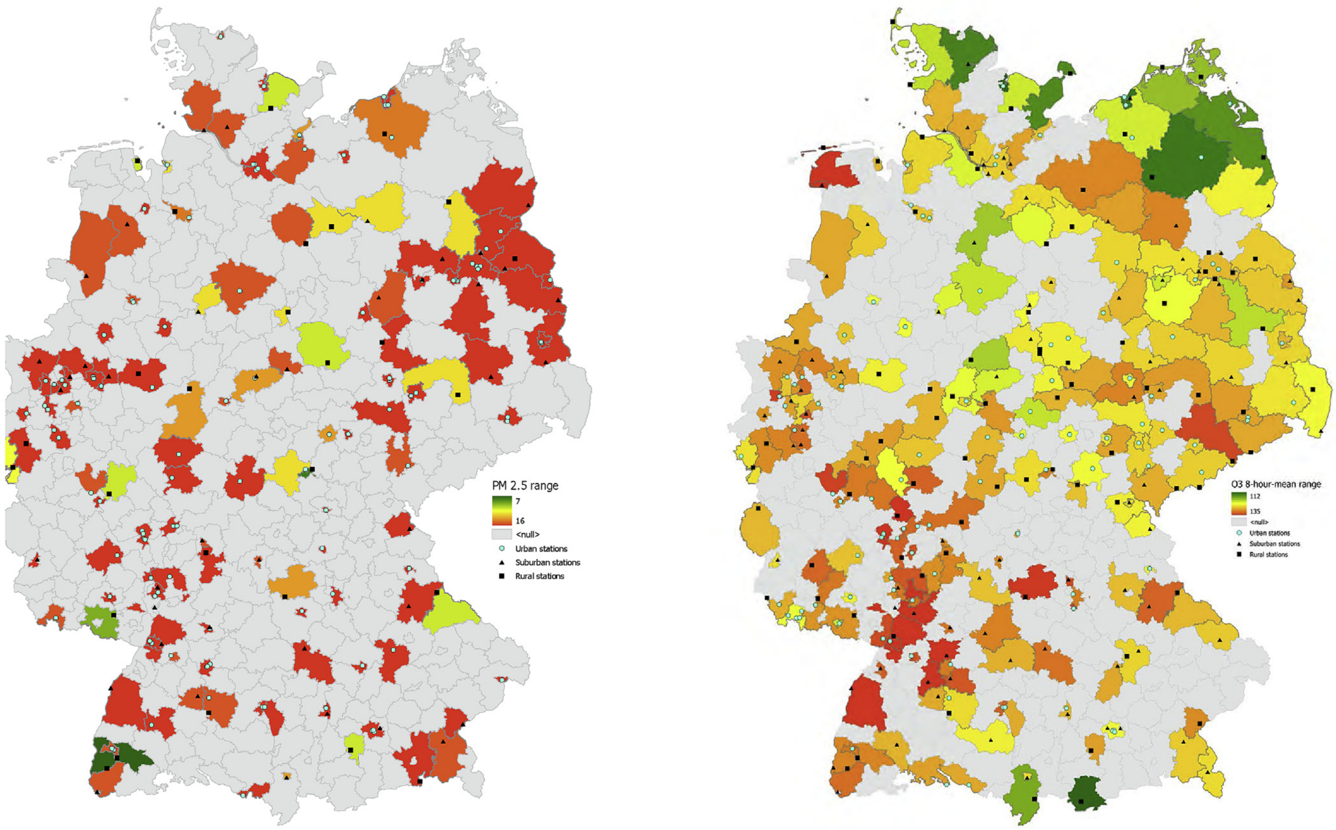
Satellite data. Gridded ground-level pollution data for NO_2 and $PM_{2.5}$ is obtained from the Atmospheric Composition Analysis Group (Geddes et al., 2015; Van Donkelaar et al., 2016). It comes at resolutions of 0.1° by 0.1° (NO_2) and 0.01° by 0.01° ($PM_{2.5}$). In order to get pollution at the district level, we took a map of German districts and calculated the mean concentration of ground-level pollution per district and year using the gridded pollution data.

$PM_{2.5}$ data itself is obtained from different satellite instruments from NASA (MODIS, MISR, and SeaWIFS), which observe backscattered solar radiation and thereby Aerosol Optical Depth (AOD, small particles floating in the atmosphere are called aerosols). These observations are then transformed into ground-level data using Chemical Transport Models (here GEOS-Chem), which simulate the geophysical relationship between AOD and $PM_{2.5}$. Afterwards, those estimations are calibrated using monitoring stations where possible. This is mostly the case in economically developed areas like the US or Europe. The resulting dataset provides annual pollution data from 1998 until 2016. An exact explanation of how the data is produced is provided by Van Donkelaar et al. (2016).

For the production of the NO_2 data, again a chemical transport model is used and the approach is generally similar to the one explained above. The dataset available goes from 1996 through 2012 and is described by Geddes et al. (2015).

[61] Note that we do not use water capacity of the topsoil and the subsoil, the depth to rock, the soil erodibility class, and the hydrological class in our main analyses. However, we also ran regressions including those variables as instruments and did not find the second stage outcomes to change significantly.

D. Further results and robustness checks



(a) Sample of $PM_{2.5}$ (b) Sample of O_3

Fig. A.2 Monitoring stations (black dots) and mean $PM_{2.5}$ (left) and O_3 (right) concentration levels in 2015.

Table A.1
R2 of bivariate regressions of population density on historical population density and soil characteristics

	NO_2	PM_{10}	$PM_{2.5}$	O_3
Population density in 1910	0.70	0.70	0.70	0.70
Ruggedness	0.07	0.05	0.06	0.03
Soil differentiation (3 categories)	0.03	0.02	0.01	0.01
Soil carbon content (4 categories)	0.24	0.22	0.28	0.21
Dominant parent material (7 categories)	0.15	0.18	0.23	0.16
Topsoil mineralogy (4 categories)	0.04	0.04	0.02	0.03
Subsoil mineralogy (3 categories)	0.03	0.02	0.03	0.02

Observations are 392 in NO_2-sample, 331 in PM_{10}-sample, 132 in $PM_{2.5}$-sample, and 253 in O_3-sample. We only kept observations in 2015 in the data.

Table A.2
IV regressions with historical working population as control variable

	NO_2		PM_{10}		$PM_{2.5}$		O_3	
	(1) IV Density 1910	(2) IV Soil	(3) IV Density 1910	(4) IV Soil	(5) IV Density 1910	(6) IV Soil	(7) IV Density 1910	(8) IV Soil
log(pop density)	0.178***	0.370***	0.0801***	0.0366	0.0452	−0.00907	−0.0911***	−0.295***
	(0.0569)	(0.0579)	(0.0273)	(0.0410)	(0.0545)	(0.0619)	(0.0329)	(0.0495)
Share employed in Ind.	0.0598		0.184**		0.376***		0.00615	
	(0.122)		(0.0826)		(0.145)		(0.0997)	
Share workers in Agr.		0.376***		−0.159**		−0.190		−0.355***
		(0.102)		(0.0727)		(0.134)		(0.0814)
Distance to CBD	−0.00514***	−0.00282**	0.000914	0.000739	0.00199	0.00126	0.00366***	0.00124
	(0.00153)	(0.00130)	(0.000881)	(0.000885)	(0.00154)	(0.00159)	(0.000997)	(0.000925)
Distance to Street	−0.0957**	−0.102***	−0.0280	−0.0410**	−0.0313	−0.0366	0.0420*	0.0513**
	(0.0399)	(0.0375)	(0.0193)	(0.0199)	(0.0430)	(0.0416)	(0.0248)	(0.0248)
Suburban	0.280***	0.288***	0.0688***	0.0685***	0.0953	0.116*	−0.143***	−0.142***
	(0.0507)	(0.0518)	(0.0247)	(0.0256)	(0.0583)	(0.0692)	(0.0363)	(0.0370)
Urban	0.474***	0.474***	0.122***	0.131***	0.123*	0.161**	−0.221***	−0.214***
	(0.0708)	(0.0720)	(0.0330)	(0.0325)	(0.0705)	(0.0821)	(0.0455)	(0.0468)
Industrial	0.0847**	0.113***	0.115***	0.129***	0.0659*	0.0895**	−0.0496	−0.107***
	(0.0382)	(0.0351)	(0.0356)	(0.0378)	(0.0362)	(0.0390)	(0.0339)	(0.0336)
Traffic	0.659***	0.649***	0.254***	0.250***	0.137***	0.123***	−0.235***	−0.246***
	(0.0407)	(0.0387)	(0.0184)	(0.0185)	(0.0436)	(0.0430)	(0.0444)	(0.0399)
N	5091	5336	4272	4485	747	769	3414	3581
R2	0.751	0.763	0.486	0.489	0.294	0.273	0.420	0.459
Districts	269	269	247	247	109	109	251	251

Standard errors in parantheses are clustered at labor market region - year (OLS) and labor market region (IV) level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.3
Station fixed effects for all pollutants including control variables

	NO_2		PM_{10}		$PM_{2.5}$		O_3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pop density)	0.290***	0.286**	0.00294	0.0241	0.336*	0.338	0.268**	0.0144
	(0.0942)	(0.119)	(0.0912)	(0.109)	(0.193)	(0.225)	(0.106)	(0.127)
Av. GDP		0.120*		0.0566		−0.108		−0.0858
		(0.0624)		(0.0548)		(0.148)		(0.0519)
Av. Income		0.238*		0.148		0.00201		−0.152
		(0.132)		(0.149)		(0.533)		(0.160)
Unemployment share		0.612*		0.452		0.528		0.654**
		(0.369)		(0.372)		(1.061)		(0.317)
Green Voters		−0.457		−1.328***		0.863		−0.243
		(0.333)		(0.493)		(1.369)		(0.364)
Env. Zone		−0.00340		−0.0167***		−0.00577		0.00836
		(0.00591)		(0.00597)		(0.00854)		(0.00697)
N	5575	4905	4648	4137	795	719	3776	3438
R2	0.094	0.107	0.010	0.028	0.075	0.102	0.040	0.043
Districts	269	269	247	247	109	109	251	251
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parantheses are clustered at district level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.4
Labor Market Regions Regressions

	NO_2		PM_{10}		$PM_{2.5}$		O_3	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
log(pop density)	0.311***	0.303***	0.0958***	0.109***	0.0718*	0.0847**	−0.233***	−0.222***
	(0.0648)	(0.0722)	(0.0248)	(0.0262)	(0.0394)	(0.0411)	(0.0439)	(0.0492)
Distance to CBD	−0.00670***	−0.00672***	0.000140	0.000188	0.000839	0.000872	0.00387***	0.00392***
	(0.00140)	(0.00141)	(0.000741)	(0.000734)	(0.00120)	(0.00117)	(0.000996)	(0.000995)
Distance to Street	−0.0971**	−0.0974**	−0.0384*	−0.0382*	−0.0347	−0.0343	0.0480**	0.0479**
	(0.0387)	(0.0384)	(0.0206)	(0.0203)	(0.0393)	(0.0383)	(0.0233)	(0.0231)
Urban	0.502***	0.503***	0.154***	0.152***	0.167***	0.163***	−0.231***	−0.232***
	(0.0602)	(0.0605)	(0.0298)	(0.0294)	(0.0591)	(0.0580)	(0.0407)	(0.0404)
Traffic	0.658***	0.658***	0.261***	0.261***	0.110***	0.111***	−0.252***	−0.251***
	(0.0364)	(0.0362)	(0.0173)	(0.0172)	(0.0413)	(0.0401)	(0.0361)	(0.0360)
N	5575	5575	4648	4648	795	795	3776	3776
*R*2	0.753	0.753	0.481	0.480	0.268	0.268	0.470	0.470
Labor Market Regions	127	127	124	124	76	76	125	125
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic		9081.8		327398.9		132.2		25302.3

Standard errors in parantheses are clustered at labor market region - year (OLS) and labor market region (IV) level.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.5
Satellite data regressions

	(1) OLS	(2) IV Density 1910	(3) IV Soil	(4) IV 1910 & Soil	(5) FE	(6) FD
Panel A: NO_2						
log(pop density)	0.374***	0.418***	0.729***	0.385***	−0.00591	−0.133
	(0.0109)	(0.0158)	(0.0237)	(0.0141)	(0.109)	(0.144)
N	4414	4161	4414	4161	4414	803
*R*2	0.247	0.249	0.059	0.252	0.364	0.338
Districts	402	402	402	402	402	402
Soil Characteristics	No	No	Yes	Yes	No	No
Panel B: $PM_{2.5}$						
log(pop density)	0.0485***	0.107***	0.0598***	0.0823***	0.421***	0.376***
	(0.00318)	(0.00451)	(0.00643)	(0.00406)	(0.0432)	(0.0802)
N	6022	5677	6022	5677	6022	803
*R*2	0.497	0.466	0.496	0.483	0.704	0.641
Districts	402	402	402	402	402	402
Soil Characteristics	No	No	Yes	Yes	No	No

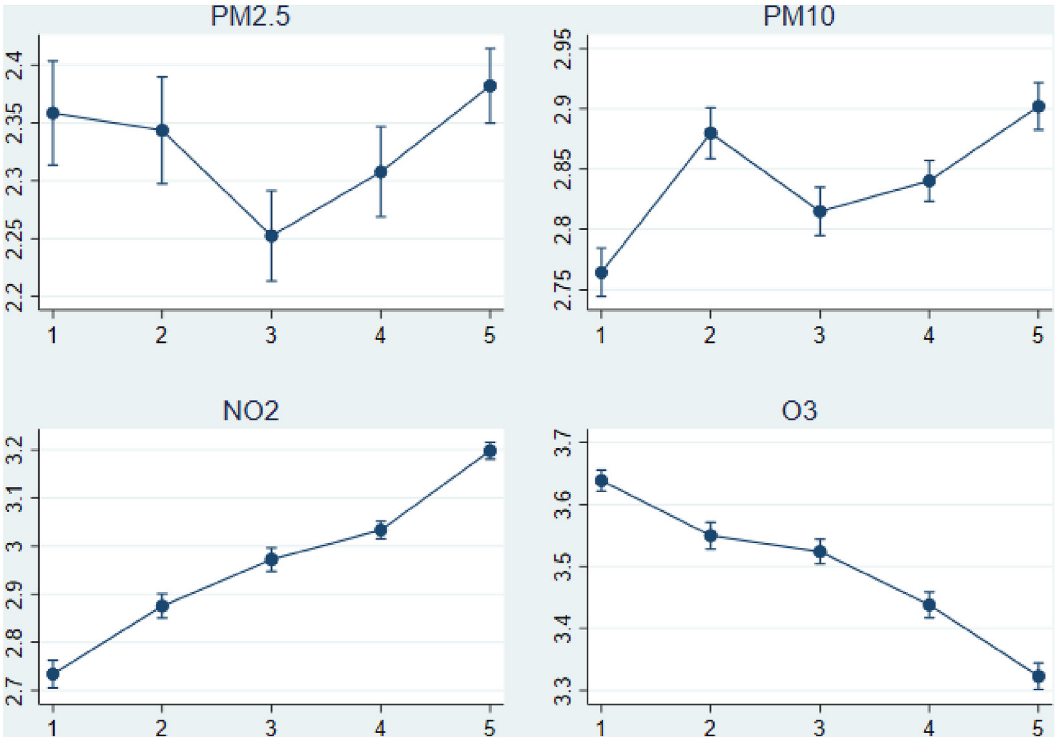


Fig. A.3 Quintiles of population density using the whole sample

Appendix E. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.regsciurbeco.2020.103596.

References

Achakulwisut, P., Brauer, M., Hystad, P., Anenberg, S.C., 2019. Global, national, and urban burdens of paediatric asthma incidence attributable to ambient NO2 pollution: estimates from global datasets. Lancet Planet. Health 3 (4), e166–e178.
Ahlfeldt, G.M., Pietrostefani, E., 2019. The economic effects of density: a synthesis. J. Urban Econ. 111, 93–107.
Auffhammer, M., Hsiang, S.M., Schlenker, W., Sobel, A., 2013. Using weather data and climate model output in economic analyses of climate change. Rev. Environ. Econ. Pol. 7 (2), 181–198.
Auffhammer, M., Kellogg, R., 2011. Clearing the air? The effects of gasoline content regulation on air quality. Am. Econ. Rev. 101 (6), 2687–2722.
Bauernschuster, S., Hener, T., Rainer, H., 2017. When labor disputes bring cities to a standstill: the impact of public transit strikes on traffic, accidents, air pollution, and health. Am. Econ. J. Econ. Pol. 9 (1), 1–37.
Blaudin de Thé, C., Carantino, B., Lafourcade, M., 2018. The Carbon 'carprint' of Suburbanization: New Evidence from French Cities. CEPR. CEPR Discussion Papers 13086.
Borck, R., 2019. Public transport and urban pollution. Reg. Sci. Urban Econ. 77, 356–366.
Borck, R., Brueckner, J.K., 2018. Optimal energy taxation in cities. J. Assoc. Environ. Resour. Econ. 5 (2), 481–516.
Borck, R., Pflüger, M., 2019. Green cities? urbanization, trade, and the environment. J. Reg. Sci. 59 (4), 743–766.
Borck, R., Schrauth, P., 2019. Urban Pollution: A Global Perspective. (Unpublished working paper).
Borck, R., Tabuchi, T., 2018. Pollution and city size: can cities be too small? J. Econ. Geogr. 19 (5), 995–1020.
Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50 (2), 317–372.
Carozzi, F., Roth, S., 2019. Dirty Density: Air Quality and the Density of American Cities. Centre for Economic Performance. CEP Discussion Papers dp.1635.
Ciccone, A., Hall, R.E., 1996. Productivity and the density of economic activity. Am. Econ. Rev. 86 (1), 54–70.
Combes, P.-P., Duranton, G., Gobillon, L., 2018. The costs of agglomeration: house and land prices in French cities. Rev. Econ. Stud. 86 (4), 1556–1589.
Combes, P.-P., Duranton, G., Gobillon, L., Roux, S., 2010. Estimating agglomeration economies with history, geology, and worker effects. In: Glaeser, E.L. (Ed.), Agglomeration Economics. University of Chicago Press, Chicago, pp. 15–66.
Combes, P.-P., Gobillon, L., 2015. The empirics of agglomeration economies. In: Duranton, G., Henderson, J.V., Strange, W.C. (Eds.), *Handbook of Regional and Urban Economics*, Volume 5 of *Handbook of Regional and Urban Economics*. Elsevier, pp. 247–348 (Chapter 5).
Costa, S., Ferreira, J., Silveira, C., Costa, C., Lopes, D., Relvas, H., Borrego, C., Roebeling, P., Miranda, A.I., Paulo Teixeira, J., 2014. Integrating health on air quality assessment — review report on health risks of two major european outdoor air pollutants: Pm and NO2. J. Toxicol. Environ. Health, Part B 17 (6), 307–340.
Duranton, G., Puga, D., 2019, December. Urban Growth and its Aggregate Implications. National Bureau of Economic Research. Working Paper 26591.
Elshout, S.v. d., Bartelds, H., Heich, H., Léger, K., 2012. Caqi Air Quality Index. Comparing Urban Air Quality across Borders. Eur. Union, Brussels, Belgium (Tech. Rep).
Ewing, R., Pendall, R., Chen, D., 2003. Measuring sprawl and its transportation impacts. Transport. Res. Rec.: J. Transport. Res. Board 1831, 175–183.
Freeman, R., Liang, W., Song, R., Timmins, C., 2019. Willingness to pay for clean air in China. J. Environ. Econ. Manag. 94, 188–216.
Geddes, J.A., Martin, R.V., Boys, B.L., van Donkelaar, A., 2015. Long-term trends worldwide in ambient NO2 concentrations inferred from satellite observations. Environ. Health Perspect. 124 (3), 281–289.
Gehrsitz, M., 2017. The effect of low emission zones on air pollution and infant health. J. Environ. Econ. Manag. 83, 121–144.
Glaeser, E.L., 1998. Are cities dying? J. Econ. Perspect. 12 (2), 139–160.
Glaeser, E.L., Kahn, M.E., 2010. The greenness of cities: carbon dioxide emissions and urban development. J. Urban Econ. 67 (3), 404–418.

Gordon, P., Cox, W., 2012. Cities in western europe and the United States: do policy differences matter? Ann. Reg. Sci. 48 (2), 565–594.
Gudipudi, R., Fluschnik, T., Ros, A.G.C., Walther, C., Kropp, J.P., 2016. City density and CO2 efficiency. Energy Pol. 91, 352–361.
Henderson, J.V., Squires, T., Storeygard, A., Weil, D., 2018. The global distribution of economic activity: nature, history, and the role of trade. Q. J. Econ. 133 (1), 357–406.
Hilber, C., Palmer, C., 2014. Urban Development and Air Pollution: Evidence from a Global Panel of Cities. vol. 175. Grantham Research Institute on Climate Change and the Environment. GRI Working Papers.
Kahn, M.E., 2010. New evidence on trends in the cost of urban agglomeration. In: Glaeser, E.L. (Ed.), Agglomeration Economics. University of Chicago Press, Chicago, pp. 339–354.
Karathodorou, N., Graham, D.J., Noland, R.B., 2010. Estimating the effect of urban density on fuel demand. Energy Econ. 32 (1), 86–92.
Koh, H.-J., Riedel, N., Böhm, T., 2013. Do governments tax agglomeration rents? J. Urban Econ. 75, 92–106.
Kosfeld, R., Werner, A., 2012. Deutsche Arbeitsmarktregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011. Raumforsch. Raumordn. 70 (1), 49–64.
Lagravinese, R., Moscone, F., Tosetti, E., Lee, H., 2014. The impact of air pollution on hospital admissions: evidence from Italy. Reg. Sci. Urban Econ. 49, 278–285.
Lamsal, L., Martin, R., Parrish, D.D., Krotkov, N.A., 2013. Scaling relationship for NO2 pollution and urban population size: a satellite perspective. Environ. Sci. Technol. 47 (14), 7855–7861.
Larson, W., Yezer, A., 2015. The energy implications of city size and density. J. Urban Econ. 90, 35–49.
Moreno-Monroy, A.I., Schiavina, M., Veneri, P., 2020. Metropolitan areas in the world. delineation and population trends. J. Urban Econ. 103242.
Newman, P.W., Kenworthy, J.R., 1989. Gasoline consumption and cities: a comparison of US cities with a global survey. J. Am. Plann. Assoc. 55 (1), 24–37.
Nivola, P.S., 1999. Fall). Are europe's cities better? Publ. Interest 73–84.
Oliveira, E.A., Andrade Jr., J.S., Makse, H.A., 2014. Large cities are less green. Sci. Rep. 4 Article number: 4235.
Panagos, P., Van Liedekerke, M., Jones, A., Montanarella, L., 2012. European soil data centre: response to european policy support and public data requirements. Land Use Pol. 29 (2), 329–338.
Pope III, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. J. Air Waste Manag. Assoc. 56 (6), 709–742.
Redding, S.J., Sturm, D.M., 2008. The costs of remoteness: evidence from German division and reunification. Am. Econ. Rev. 98 (5), 1766–1797.
Sarzynski, A., 2012. Bigger is not always better: a comparative analysis of cities and their air pollution impact. Urban Stud. 49 (14), 3121–3138.
Schindler, M., Caruso, G., Picard, P., 2017. Equilibrium and first-best city with endogenous exposure to local air pollution from traffic. Reg. Sci. Urban Econ. 62, 12–23.
Van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., Winker, D.M., 2016. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. 50 (7), 3762–3772.
WHO Regional Office for Europe and OECD, 2015. Economic Cost of the Health Impact of Air Pollution in Europe: Clean Air, Health and Wealth. WHO Regional Office for Europe, Copenhagen.
Wolff, H., 2014. Keep your clunker in the suburb: low-emission zones and adoption of green vehicles. Econ. J. 124 (578), F481–F512.
World Health Organization, 2003. Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide. WHO Regional Office for Europe, Copenhagen.
World Health Organization, UNAIDS, etal, 2006. Air Quality Guidelines: Global Update 2005. World Health Organization.
Zhou, Y., Levy, J.I., Evans, J.S., Hammitt, J.K., 2006. The influence of geographic location on population exposure to emissions from power plants throughout China. Environ. Int. 32 (3), 365–373.