CESifo WORKING PAPERS

7629
2019

April 2019

Population density and urban air quality

Rainald Borck, Philipp Schrauth

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7629*
Category 9: Resource and Environment Economics

Population density and urban air quality

Abstract

We use panel data from Germany to analyze the effect of population density on urban air pollution (nitrogen oxides, particulate matter and ozone). 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. Our preferred estimates imply that a one-standard deviation increase in population density increases air pollution by 3-12%.

JEL-Codes: Q530, R120.

Keywords: population density, air pollution.

Rainald Borck
University of Potsdam / Germany
rborck@uni-potsdam.de

Philipp Schrauth
University of Potsdam / Germany
schrauth@uni-potsdam.de

April 18, 2019
We thank Gabriel Ahlfeldt as well as participants at the PRSE seminar and the workshop “Cities and the environment” in Potsdam and the UEA meetings in Düsseldorf for helpful comments. We are grateful to the German Research Foundation (DFG) for financial support.

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 these costs than about agglomeration economies.[1] Kahn (2010) documents that in the US, larger cities have longer commute times, higher pollution levels and higher crime rates. In this paper, we 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 developing countries, where urban air pollution has fallen over the last decades, high pollution levels keep occurring. German cities have been subject to a variety of legal proceedings against transgressions of pollution thresholds, and the same is true of other European cities. 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.[2] 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, 2015). The European Environment Agency estimates that in Germany, particulate matter ($PM_{2.5}$) caused 66,000 premature deaths in 2013.[3] 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 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

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

[2] While evidence of the health effects of NO_2 is rather scarce, there is wide evidence on the health effects of particulate matter (see e.g. Pope III and Dockery (2006) for a summary).

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

fuels for heating or energy production.[4] 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 e.g. in paints or in gasoline exhausts). 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. NO_2 is a toxic gas, which is damaging to human health at explicitly high levels of more than 200 $\mu g/m^3$ in the short run. 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).

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}, $PM_{2.5}$, and O_3) for German cities, using rich panel data from 2002 to 2015. We start by presenting OLS estimates. However, these may be biased due to omitted variables or reverse causality, so we also estimate fixed effects (and long difference) 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, a one-standard deviation increase in population density increases PM_{10} concentration by about 3% and NO_2 concentration by about 12%. Thus, this is one of the first papers to estimate the causal effect of population density on pollution.

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.

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

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 (2015) 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) also simulate the energy implications of city size and density. They find that per-capita energy use is relatively invariant to city size when growth is driven by wages but falls modestly with growth induced by better amenities.

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 residents of denser cities consume less energy per capita (Glaeser and Kahn, 2010; Blaudin de Thé et al., 2018). The reason is that, as analyzed by the theoretical literature, residents of densely populated cities consume less fuel due to the availability of public transport systems and 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 pollution (Bauernschuster et al., 2017; Borck, 2017). Gudipudi et al. (2016) use U.S. data from 2000 and find that per capita CO_2 emissions decrease with city size. Oliveira et al. (2014) use the same dataset and the same method, but a slightly different definition of cities. They find per capita CO_2 emissions increase with city size. 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. Bart (2010) examines urban sprawl in European cities and argues that urban sprawl leads to a strong increase in CO2 emissions caused by transport.

Another set of papers examines the effect of population size and other explanatory factors on air quality, mostly particulate matters, sulphur dioxide and nitrogen oxides. In an early study, Glaeser (1998) finds that particulate matter levels increase with city size and calculates the costs of a city increasing in size from 500,000 to 5,000,000 people to lie between \$38 and \$185 per person annually. Lamsal et al. (2013) examine cities on different continents in 2005 and find significant positive relationships between population and NO_X emissions on all continents. Sarzynski (2012) uses OLS regressions on a sample of cities worldwide in 2005 and finds that the number of inhabitants in a city is significantly positively correlated with NO_X, while urban density has a significant negative effect on emissions. Ewing et al. (2003) find a negative correlation between density and mean annual NO_X and VOC emissions. Stone (2008) finds a positive correlation between urban sprawl and the number of exceedances of yearly ozone standards. Ahlfeldt and Pietrostefani (2019) use a cross-section of 343 cities worldwide to study the effect of population density on particulate ($PM_{2.5}$) concentration and find an elasticity of around 0.125. Hilber and Palmer (2014) analyze pollution using panel data from 75 metropolitan areas in 45 OECD and non-OECD countries. Using fixed effects and controlling for a rich set of explanatory variables, they find that an increase of population density leads to negative effects for PM_{10} and NO_2 concentrations and therefore an improvement in air quality. Carozzi and Roth (2019) study the effect of population density on $PM_{2.5}$-concentration in the US and find a statistically significant positive elasticity of 0.13, which is slightly larger than the effect we find.

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 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.[5] 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. Finally, our paper contains data on other pollutants as well (PM_{10}, NO_2, and O_3), 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.

[5] To mitigate the latter problem, we will include some station characteristics, such as distance to city center, station type, 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.

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 (2015), Borck and Tabuchi (2018) and Larson and Yezer (2015). More details are in Appendix A. Consider a circular monocentric city with N workers who commute to the CBD for work. Households have utility $v(c, q)$ over consumption, c, and square meters of housing floor space, q. 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. 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, $r(\bar{x})$ must equal the opportunity cost of land r_A. 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.[6] 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.[7] Pollution concentration in a city is the sum of total emissions divided by land area.[8]

Suppose that city population rises. Then, the city expands spatially, and population density rises. Average commuting distances increase, which increases traffic emissions. On the other hand, residents will reside in smaller dwellings. This effect will tend to decrease per capita energy use, while total housing supply and thus total residential energy use increases. As a result, total pollution will rise. Numerical simulations show that pollution concentration also rises (see Appendix A).

We can also show (using numerical simulation) that making the city denser by restricting its spatial extension (through increasing agricultural land rents) increases pollution concentration. This happens even though the denser city has less total commuting and lower residential energy use, as competition for central land increases and makes

[6] Borck and Pflüger (2015) in addition consider emissions from industrial and agricultural production, and intercity goods transport, while Borck (2017) considers the effect of modal choice between public and private transport.

[7] Borck and Brueckner (2018) propose a model where instead emissions from residential energy use are proportional to the building's surface area, which leads to scale economies in residential energy use, since taller buildings have a lower surface per unit of floor space. Note also that we abstract from congestion, see e.g. Larson and Yezer (2015).

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

dwellings smaller. But the city area decreases even more, so pollution concentration rises. This model therefore predicts that the concentration of pollution from transport and residential energy rises with density.[9]

However, this model ignores some possible countervailing forces. First, due to their larger density, bigger cities tend to have a more extensive supply of public transit. Since transit typically produces lower emissions per person kilometer than automobiles, this would tend to decrease traffic emissions, all else equal. And 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). All in all, concentration in bigger cities might conceivably fall with population density. In the next sections, we will examine the empirical relation between density and air pollution.

4 Data and estimation

4.1 Data

We use administrative panel data from Germany for the period 2002 - 2015. While we have hourly data on all emissions monitor stations in Germany, our regional data, in particular population density, is available on a yearly basis for the roughly 400 German counties (Landkreise).

4.1.1 Emissions and weather 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. 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).

[9] Note that, unlike Glaeser and Kahn (2010) we focus on pollution concentration rather than per household emissions.

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 and each hour respectively. 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 the next subsection.

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. In order to rule out the possibility that results are driven by cyclical 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 that is not related to population density.

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. Note, however, that some of the thresholds defined by political institutions lie somewhat higher than the ones suggested by the WHO.

The World Health Organization (WHO) has published guidelines for pollution concentration levels based on potential health threats (Tab. 1). For PM_{10}, these are 20 $\mu g/m^3$ for the annual mean concentration and 50 $\mu g/m^3$ for the 24-hour 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-hour 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 one-hour mean value of 200 $\mu g/m^3$.[10] Even though the WHO stresses the importance of thresholds for health impacts, to our knowledge there is no

[10]The guidelines set by 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-hour 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

Table 1: WHO thresholds of NO_2 and PM_{10}

	NO_2	PM_{10}	$PM_{2.5}$
Annual mean	$40\mu g/m^3$	$20\mu g/m^3$	$10\mu g/m^3$
24-hour mean		$50\mu g/m^3$	$25\mu g/m^3$
Hourly mean	$200\mu g/m^3$		

research which explicitly analyzes these thresholds. The literature on potentially non-linear effects of pollution on health – e.g. when crossing a particular threshold – is scarce and the issue has recently been debated in the public. Still, we think that analyzing threshold transgressions is interesting, because of potentially increasing health effects and the focus of public attention on these thresholds.

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.[11] 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.[12] The German Meteorological Service (DWD in German) provides free access to the data of its various weather and precipitation stations. This allows us to get hourly data on temperature, air pressure, rainfall, snowfall, sunshine, and wind. 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 B.

4.1.2 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 moni-

the 24-hour-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 $25\mu g/m^3$.

[11] 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 are available upon request).

[12] As Auffhammer and Kellogg (2011) note, ozone creation needs a certain amount of NO_2 and of other volatile organic compounds. 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.

toring station. We can control for a set of station-specific factors such as the distance to the central business district (CBD),[13] whether the station lies in an urban or a suburban area and the station type (traffic, background, or industry, see Section 4.1.1).

In Germany, over the course of the past 15 years, many cities introduced low emission zones (*Umweltzonen*). Those zones were established in order to lower high levels of particulate matter by restricting city access to cars that have particle filters.[14] Using maps, we assign to each monitoring station an indicator for whether or not it lies in a low emission zone. 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. However, if the implementation of low emission zones is a political reaction to high pollution caused by higher density, then the actual effect we want to measure would suffer from downward bias. This is why we will include this control variable in robustness regressions only.

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 share of voters for the Green Party, in order to control for the potential sorting of 'green' households into cities.

For reasons explained below, our main results are from regressions that control only for station type, distance to the CBD and urban status. However, to check for the robustness of our results, we include the other control variables in sensitivity checks.

Even though we have information on whether a monitoring station is measuring traffic or background pollution, outcomes especially of background stations could be driven by their distance to big street axes.[15] Therefore, in additional specifications, we control for the distance of a measuring station to the next major road (*Bundesstraße* or a street of similar size).

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.[16] Since burning coal is a major source of air pollution, this might take out some variation that is caused by the presence of coal mines.

[13]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.

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

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

[16]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.

4.1.3 Descriptives

Table 2: Descriptives

	NO2	PM10	PM2.5	O3
Overall Stations	623	531	146	409
Background	360	290	79	340
Industrial	43	44	15	26
Traffic	220	197	52	43
Districts	269	247	108	251
Urban Districts	88	85	51	72
Labor market regions	128	125	77	126
Stations in LEZ	92	94	26	34
Whole Sample				
Mean Pollution	28.11	22.52	14.43	46.90
S.D. Pollution	15.17	5.767	3.026	9.934
Mean Popdensity	2593.0	2531.1	2384.6	2254.9
S.D. popdensity	1341.6	1332.5	1328.9	1266.5
Sample in 2015				
Mean Pollution	25.87	19.00	12.93	49.72
S.D. Pollution	14.72	3.824	1.716	8.583
Mean Popdensity	2569.8	2477.5	2326.0	2215.0
S.D. popdensity	1419.3	1362.8	1323.3	1385.4

Table 2 shows monitoring stations in our sample in 2015 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, while $PM_{2.5}$ is measured by less than 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 over 404 districts including urban districts.[17] Instead of 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

[17] 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.

in about half of the labor market regions. Furthermore, the variation in pollution levels for NO_2 is relatively high compared to particulate matter.

For a first visual impression about the relation between population density and air pollution, Figure 1 shows quintiles of NO_2 and PM_{10} concentrations, while Figure 2 shows quintiles of population densities in 1910[18] — which we will use as an instrument in the subsequent parts of the paper — and in 2015 for districts. 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 paper.[19]

Figure 3 shows the variation of pollution over time and across pollutants. It depicts the distribution of emission levels by rank of the station for the years 2003 and 2015, for both NO_2 and PM_{10}. The horizontal red lines show annual threshold levels for NO_2 and PM_{10}.[20] For PM_{10}, we see a clear downward trend in emission levels. While in 2003 (green dots) most of the stations exceeded the WHO thresholds and some also the EU threshold (which is the same for the annual mean of PM_{10} and NO_2), in 2015 only the WHO threshold was exceeded by a few stations while all stations fell below the EU threshold. For NO_2, the picture is different. While there is a slight reduction of emissions at the bottom end of the distribution, threshold violations have not actually fallen. The widespread violations of threshold levels have led to treaty violation proceedings against the German government as well as individual lawsuits against city governments. This is also the driving force behind the recent political discussion in Germany to ban Diesel cars from cities.

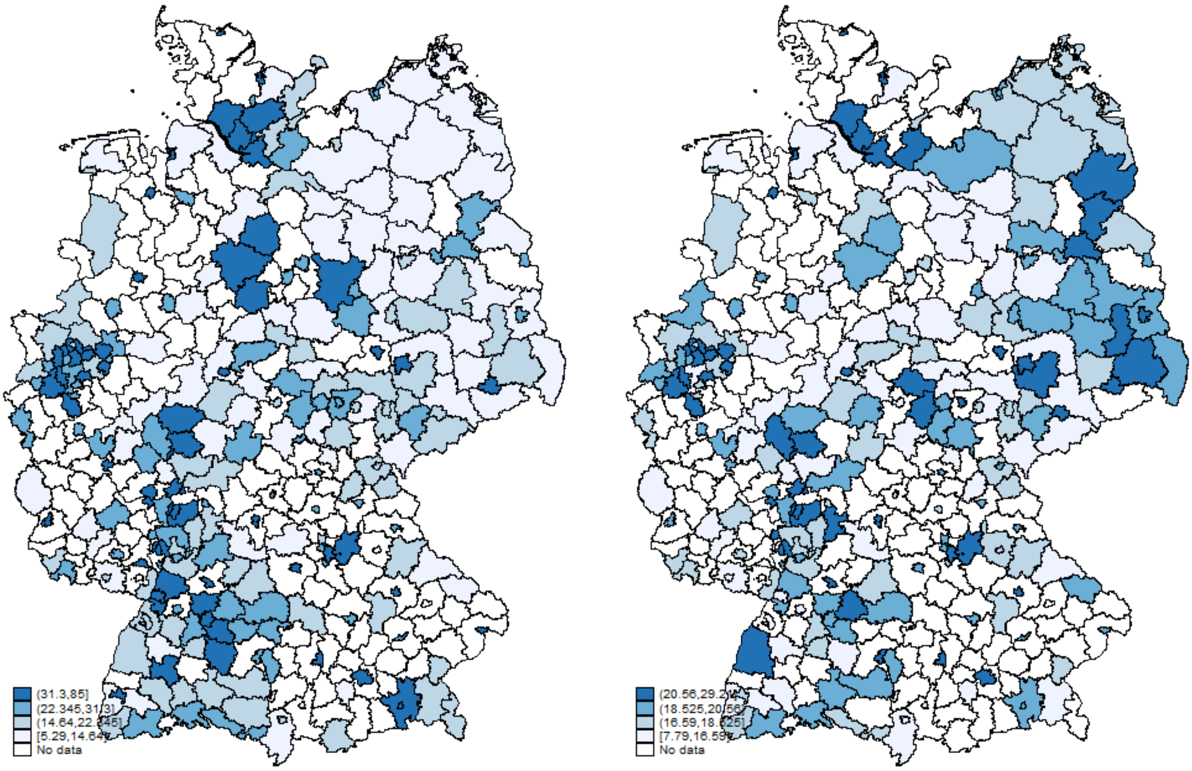
Figure 4 divides the sample into low and high density areas at the median. Except for $PM_{2.5}$(not shown), all pollutants followed a similar trend between the density groups, but at different levels. Especially PM_{10} experienced an overall decline over the time period observed. $PM_{2.5}$, which is more harmful than PM_{10}, increased slightly in low-density

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

[19]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.

[20]See Section 4.1.1 above for a description of the thresholds.

areas and experienced a decline in high-density areas, such that absolute levels in 2015 were almost identical in the two types of districts.



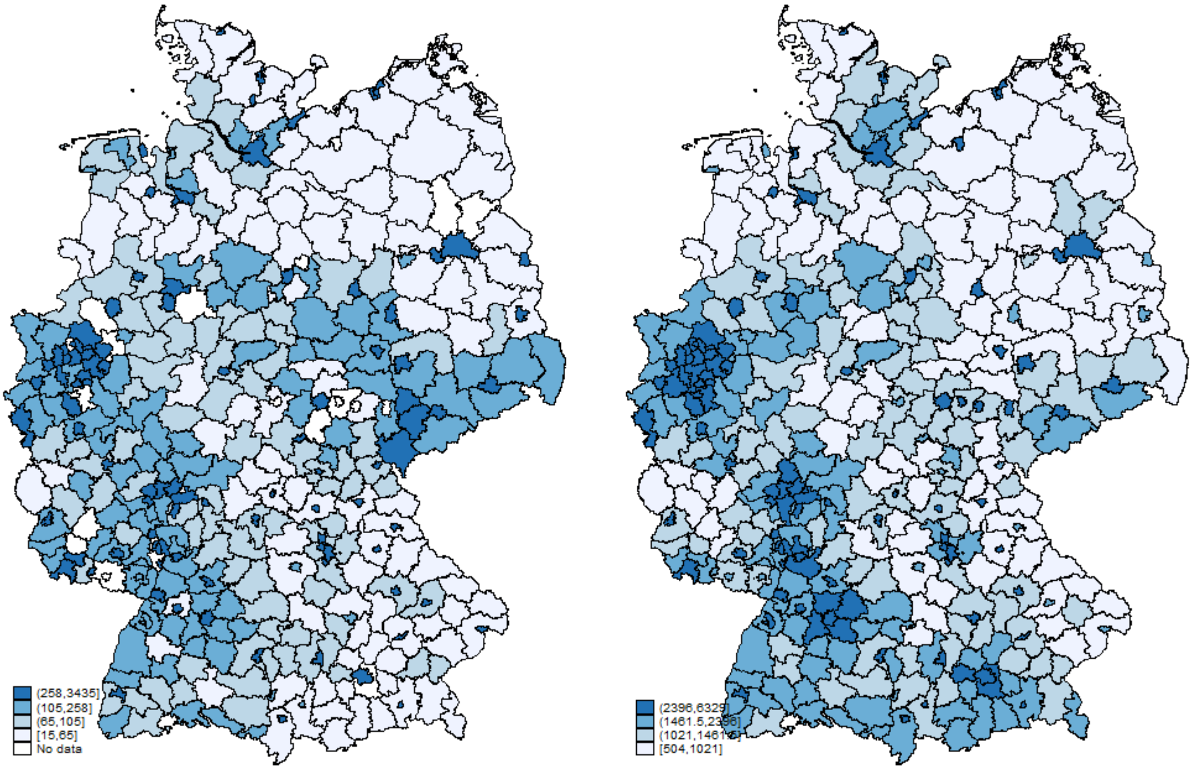
(a) NO_2 (b) PM_{10}

Figure 1: Mean NO_2 (left) and PM_{10} (right) concentration levels

4.2 Estimation

4.2.1 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 (hour of day, day of week and month dummies). Following Auffhammer et al. (2013), the extensive set of weather and weather-interaction variables includes the hourly level of precipitation, sunshine, wind-speed, cloudiness, and temperature at weather stations, as well as quadratic terms for all of these variables and a cubic temperature variable. We also interact temperature with wind. We include as further controls an indicator for working days (Monday until Friday), 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



(a) 1910 (b) 2015

Figure 2: Development of population densities over time

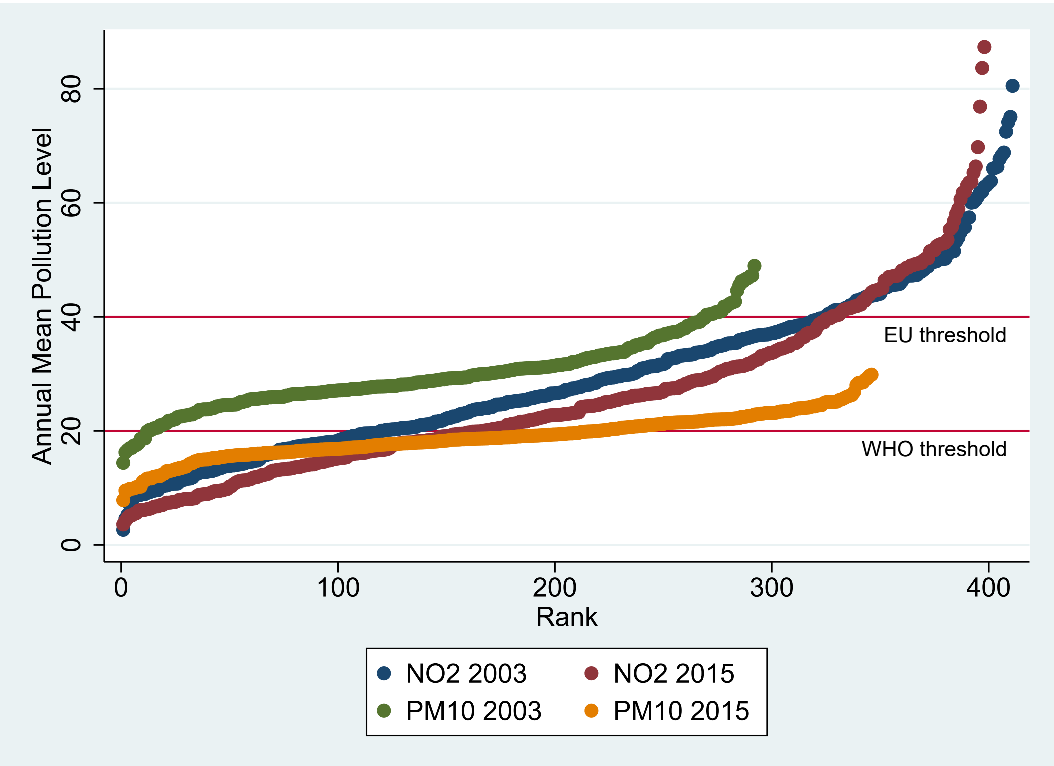


Figure 3: Correlation of ranks and pollution

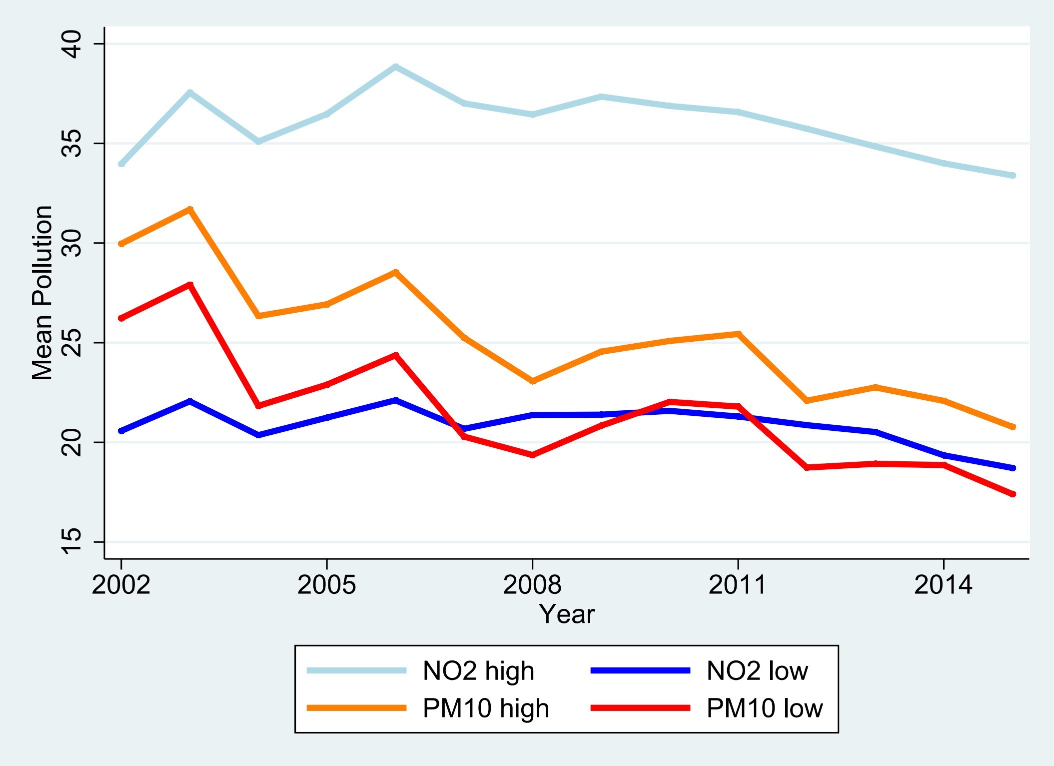


Figure 4: Pollution in high and low density areas over time

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 estimated residuals and start with a simple OLS model and regress hourly pollution outcomes on a set of control variables. Our main regressor of interest is population density, which is available in yearly intervals at the district level. Our first regression equation is

$$\ln(Y_{it}) = \beta + \rho \ln(D_{it}) + \gamma \mathbf{X_{it}} + \alpha_t + \epsilon_{it}, \tag{1}$$

where Y_{it} denotes the residual concentration level (for a particular pollutant) in year t at station i. Like Henderson (1996), 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 $\mathbf{X}$. Those are attributes of the monitoring station like the station area (urban, suburban or rural) and station type (background, traffic, or industrial). 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 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_{it}. 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. Apart from that we calculated the distance of a measuring station to the closest coal-fired power plant.

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.

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 many 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 large districts with large energy demand. Therefore, we choose to present our basic regressions with controls only for the urban/suburban/rural indicator, station type and distance to the CBD. As a sensitivity check, we analyze in Appendix C how our results change when we successively add controls.

OLS estimates would be unbiased and consistent as long as 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.[21] Therefore, we also estimate long difference and fixed effects regressions of the form

$$\Delta \ln(Y_{it}) = \rho \Delta \ln(D_{it}) + \gamma \Delta \mathbf{X_{it}} + \Delta \alpha_t + \Delta \tilde{\epsilon}_{it}, \tag{2}$$

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 C). 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 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 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 variation of density and pollution within only is low, fixed effects estimates may suffer from imprecise estimates. Therefore, we also estimate instrumental variables (IV) regressions:

$$\ln(D_{it}) = \theta + B_1 W_{it} + B_2 X_{it} + B_3 Z_{it} + \eta_{it} \tag{3}$$

$$\ln(Y_{it}) = \vartheta + \rho \widehat{\ln(D_{it})} + A_1 W_{it} + A_2 X_{it} + \hat{\epsilon}_{it} \tag{4}$$

Here, in the first stage regression (3), density is regressed on one or more instrumental variable(s). 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 from 1910.[22] Historical

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

[22]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.

population data are relevant, since urban population tends to be strongly persistent over time. The exclusion restriction requires that historical density is 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, so that correlation between historical density and current pollution seems unlikely. However, in order to control for the possibility that industry structures persist over time, we also control for the share of workers in industry and crafts in 1925.[23] Appendix B describes how we constructed the corresponding variables.

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.[24] 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 include the same 12 variables from the European Soil Database (ESDB) used by 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 water capacity of the topsoil and the subsoil, depth to rock, soil erodibility class, topsoil organic carbon content, soil profile differentiation, and the hydrological class, which describes the circulation and retention of underground water. The last variable we use is the ruggedness of a district, which is the difference of 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. More detail on these data

[23]Unfortunately, data for earlier years is not available for this variable on district level.

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

can be found in Appendix B. In all of our instrumental variable regressions we cluster standard errors at labor market region level. Since all our instruments are time invariant, we do not cluster at year level here.

4.2.2 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 4.2.1 for annual thresholds, 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-hour 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 shown in Table 1. The number of days we choose are motivated by the number of days with thresholds 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 there is no short-term threshold for $PM_{2.5}$ in the EU, there are also no allowed daily violations. Thus, we take the same value of exceedances for this pollutant as for PM_{10}. Furthermore, we look at the probability of violations on a certain number of days just below the EU allowances (15 and 10 days for NO_2 30 and 25 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 if of interest.

We then use a 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.

5 Results

5.1 Basic results

OLS regressions. We present our basic cross-sectional OLS results in columns (1) and (5) of Tab. 3 and Tab. 4. The tables present coefficients for our parameter of

Table 3: OLS and IV regressions for NO_2 and PM_{10}

	NO_2				PM_{10}			
	(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)	0.281***	0.184***	0.284***	0.244***	0.0747***	0.102***	0.0596*	0.0718***
	(0.0137)	(0.0518)	(0.0570)	(0.0519)	(0.00691)	(0.0244)	(0.0362)	(0.0245)
Distance to CBD	-0.00384***	-0.00556***	-0.00380***	-0.00461***	0.000692**	0.000871	0.000457	0.000406
	(0.000466)	(0.00158)	(0.00145)	(0.00153)	(0.000284)	(0.000858)	(0.000865)	(0.000818)
Suburban	0.345***	0.340***	0.344***	0.335***	0.103***	0.0923***	0.104***	0.0933***
	(0.0152)	(0.0462)	(0.0451)	(0.0462)	(0.00882)	(0.0233)	(0.0246)	(0.0236)
Urban	0.513***	0.541***	0.512***	0.521***	0.168***	0.142***	0.174***	0.153***
	(0.0185)	(0.0579)	(0.0589)	(0.0600)	(0.01000)	(0.0297)	(0.0306)	(0.0287)
Industrial	0.0888***	0.0886**	0.0888**	0.0877**	0.136***	0.131***	0.136***	0.131***
	(0.0133)	(0.0414)	(0.0357)	(0.0379)	(0.0126)	(0.0362)	(0.0362)	(0.0360)
Traffic	0.661***	0.669***	0.661***	0.665***	0.260***	0.261***	0.260***	0.262***
	(0.0126)	(0.0409)	(0.0392)	(0.0407)	(0.00670)	(0.0182)	(0.0177)	(0.0182)
N	5575	5301	5575	5301	4648	4407	4648	4407
R^2	0.749	0.744	0.749	0.749	0.469	0.459	0.468	0.463
Districts	269	269	269	269	247	247	247	247
Soil Characteristics	No	No	Yes	Yes	No	No	Yes	Yes

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 $PM_{2.5}$ 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)	0.0315**	0.0710	0.0351	0.0208	-0.176***	-0.0863***	-0.187***	-0.133***
	(0.0160)	(0.0558)	(0.0566)	(0.0458)	(0.00938)	(0.0308)	(0.0434)	(0.0315)
Distance to CBD	0.00105	0.00134	0.00111	0.000585	0.00274***	0.00404***	0.00257***	0.00339***
	(0.000641)	(0.00154)	(0.00147)	(0.00135)	(0.000297)	(0.000988)	(0.000958)	(0.000945)
Suburban	0.168***	0.160***	0.167***	0.165***	-0.167***	-0.167***	-0.166***	-0.162***
	(0.0226)	(0.0502)	(0.0474)	(0.0511)	(0.0111)	(0.0344)	(0.0339)	(0.0335)
Urban	0.209***	0.177***	0.207***	0.209***	-0.225***	-0.247***	-0.221***	-0.233***
	(0.0244)	(0.0549)	(0.0571)	(0.0537)	(0.0119)	(0.0380)	(0.0397)	(0.0378)
Industrial	0.0699***	0.0744*	0.0697*	0.0760*	-0.0688***	-0.0430	-0.0709**	-0.0522*
	(0.0202)	(0.0410)	(0.0396)	(0.0405)	(0.0115)	(0.0342)	(0.0295)	(0.0290)
Traffic	0.115***	0.115***	0.116***	0.110***	-0.233***	-0.239***	-0.233***	-0.240***
	(0.0162)	(0.0391)	(0.0386)	(0.0400)	(0.0162)	(0.0399)	(0.0374)	(0.0384)
N	795	758	795	758	3776	3588	3776	3588
R^2	0.246	0.223	0.246	0.228	0.437	0.416	0.436	0.431
Districts	109	109	109	109	251	251	251	251
Soil Characteristics	No	No	Yes	Yes	No	No	Yes	Yes

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

interest, log of population density, as well as our basic controls (distance to the CBD, 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 Tab. 3, 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 1.13 $\mu g/\text{m}^3$, or 12.4 percent of the mean concentration.

For PM_{10}, we find a smaller elasticity of 0.075, which is significant at 1% (column (5) of Tab. 3). A one standard deviation increase in population density increases the PM_{10} concentration by 0.057 $\mu g/\text{m}^3$ or 3.2%. For $PM_{2.5}$, the estimated elasticity is 0.03 which is significant at the 5% level (column (1) of Tab. 4). Note, however, that a net of monitoring stations for this pollutant has only recently been established, so the power of the estimates for $PM_{2.5}$ is small.

Ozone seems to be negatively correlated with population density. This is probably due to the fact that the chemical prerequisites for ozone formation are more favorable outside large cities.[25]

IV regressions. We now turn to the IV regression results. Tab. B.1 shows the results of the first stage regressions using historical population densities, the soil IV and both sets of instruments together. The F-statistics imply that our instruments are strong.

While judging from the overidentification tests alone, we cannot determine which instrument is more likely to be exogenous, we do think that the exclusion restriction is *prima facie* more credible for the soil characteristics, as argued above. While we are somewhat more confident about the soil IV, we keep both instruments in the regressions for comparison.

The IV results have the same sign and are roughly similar to the OLS results in magnitude. However, there are slightly different patterns when using historic density, using soil characteristics or both instruments jointly. In the case of the historical instruments, the changes in the point estimates are larger than for the soil IV. The point estimate decreases in our NO_2 regressions to 0.184 and increases to 0.102 in the PM_{10} regressions, but remains highly significant in both cases.

When we use the soil characteristics as instruments, the density coefficient increases slightly in the NO_2 regressions and decreases slightly in the PM_{10}regressions (Tab. 3 columns (3) and (7)). For $PM_{2.5}$, the density coefficient is of similar magnitude as the OLS one using the soil IV, but is less precisely estimated (Tab. 4 column (3)).[26] For O_3, the coefficient with the soil IV is slightly larger in absolute value than the OLS coefficient (Tab. 4 column (9)). According to the estimates, the density elasticity is 0.284 for NO_2, 0.060 for PM_{10}, and 0.035 for $PM_{2.5}$. Thus, a one standard deviation increase of population density increases the NO_2 concentration by 12.55% at the mean, and the PM_{10} concentration by 2.51%. In general, the IV results using soil characteristics as instruments are fairly similar to the OLS results. In summary, it seems that the bias from omitted variables in OLS regressions is small, a point also made by Combes et al. (2010).

[25] This is because nitrogen monoxide (NO), which is contained in car exhaust fumes, reacts with ozone to NO_2. Ozone is therefore split into O_2 and NO_2 such that ozone pollution in city centres 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`.

[26] Note, however, that due to the small sample size, statistical power of our $PM_{2.5}$ regressions in general is rather low (see Fig. B.1).

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)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE All years	LD 2002-15	FE All years	LD 2002-15	FE All years	LD 2002-15	FE All years	LD 2002-15
log(pop density)	0.347***	0.379**	-0.0225	-0.101	0.311	0.624	0.242	0.340
	(0.132)	(0.175)	(0.0907)	(0.157)	(0.224)	(1.645)	(0.156)	(0.236)
Distance to CBD	-0.0109***	-0.00951***	-0.00344**	-0.00226	-0.000518	-0.00115	0.00539***	0.00292
	(0.00220)	(0.00259)	(0.00146)	(0.00233)	(0.00396)	(0.00599)	(0.00153)	(0.00217)
Suburban	0.377***	0.355***	0.0913***	0.0152	0.0208	0.0547	-0.164***	-0.155**
	(0.0652)	(0.0902)	(0.0337)	(0.0565)	(0.0838)	(0.178)	(0.0557)	(0.0743)
Urban	0.614***	0.590***	0.179***	0.119**	0.117	0.166	-0.246***	-0.263***
	(0.0453)	(0.0472)	(0.0367)	(0.0548)	(0.0939)	(0.174)	(0.0609)	(0.0751)
Industrial	0.136***	0.158***	0.127***	0.0473	0.0554	-0.0115	-0.0820*	-0.0773
	(0.0395)	(0.0416)	(0.0420)	(0.0513)	(0.0421)	(0.0892)	(0.0448)	(0.0783)
Traffic	0.726***	0.680***	0.275***	0.262***	0.261***	0.231**	-0.211***	-0.225**
	(0.0335)	(0.0405)	(0.0152)	(0.0239)	(0.0311)	(0.101)	(0.0447)	(0.0862)
N	5575	781	4648	545	795	135	3776	549
R^2	0.895	0.894	0.761	0.804	0.794	0.932	0.823	0.833
Districts	269	258	247	235	109	105	251	248

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

Fixed effects and long differences. Fixed effects regressions may be a proper response to 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. Therefore, we present long-difference estimates for the years 2002-2015 as well as districts 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 the case 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 0.379 for the long difference regressions and 0.347 when looking at district fixed effects, and both coefficients are highly significant.

We also ran station fixed effects regressions, which are presented in table B.3. The magnitude of the coefficient in the NO_2 regressions is very similar to the one in the OLS and IV regressions.

5.2 Sensitivity.

We perform a number of robustness checks to see how sensitive the results are to various specifications. Results are in Appendix C.

First, we check how sensitive the results are to the inclusion of controls. We start with population density as the only explanatory variable and successively add further control variables to the OLS regressions. Results are shown in Tables B.4 to B.7. Looking at the results for NO_2, we see that including year fixed effects hardly changes the results (column 2). Adding station-specific control variables (urban/suburban, distance to CBD, station type; column 3) cuts the coefficient on population density in half. The presence of a coal-fired power plant may be a driving element 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 the same (column 4 of Tab. B.4). In column (5) we replace the indicator variable with a measure of the distance to the closest coal-firing power station. This variable may be better able to capture possible spillovers from coal firing power plants. Stations close to such coal-fired power plants might be more affected by its polluting output compared to monitoring stations further 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, while adding the vote share of green party voters, an indicator for low emission zone and distance to the next major street (column 7) lower the coefficient only to a minor degree. Column (8) finally adds an indicator for the state (*Bundesland*) in order to control for state-specific policies. The outcomes remain rather stable in size 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 our PM_{10} outcomes. Here, however, adding economic indicators and variables like green party voters, environmental zone indicator and distance to street increases the coefficient (columns (6) and (7) in Tab. B.5), while adding state indicators reduces it again; the coefficient remains highly significant throughout all of the specifications.

For $PM_{2.5}$ (Tab. B.6), 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 (column 4, 5 and 8).[27] However, looking at the sample distribution of our data in Figure B.1, 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 Hamburg,

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

Berlin, Munich and large parts of North Rhine-Westphalia 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.

For O_3, the picture is similar to the NO_2 outcomes (Tab. B.7): as soon as we add control variables, the coefficient is cut in half but remains highly significant throughout our specifications. In contrast to the other pollutants, however, the density coefficient is negative, so high density cities seem less affected by O_3.

In Tables B.8 to B.11 we redo this exercise with our IV estimates, but now we only successively add a larger subset of controls at a time. The picture for the IV outcomes is very similar to the OLS ones.

An interesting issue is whether the effects of population density are heterogeneous between different definitions of cities and areas of interest. To investigate this issue, we restrict the sample to district-free cities (*kreisfreie Städte*). The administrative boundaries of such districts include only one city (with the smallest city in our sample having 35.000 inhabitants), whereas 'regular' districts typically include several smaller cities and towns. As Tables B.12 and B.13 show, the coefficients on population density almost double in size. Thus, it seems that the effect of population density on air pollution may be more pronounced in urban and more populous districts.

Next, we consider a different geographical definition of "cities". Districts, whether regular ones or district-free cities, suffer from the disadvantage that they are confined within administrative boundaries. This obviously makes for an arbitrary city definition. An alternative definition is based on economic relations between cities, usually measured by commuting. We therefore rerun our basic regressions for German labor market regions (*Arbeitsmarktregionen*) as defined by Kosfeld and Werner (2012). There are 141 labor market regions defined by significant commuting flows between cities within the region. Of these regions, up to 128 are covered by our analyses. Results are shown in Tab. B.14 and B.15. The results are very close to the estimates for districts. For $PM_{2.5}$, the IV results turn significant (using the historical IV), while the soil IV returns insignificant coefficients for both particulates. This seems to be caused in part by the lower number of labor market regions compared to districts, which reduces the variation in density and hence also leads to lower precision of the estimates.

Our population density variable so far has been defined as the 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 Tab. B.16–B.19. 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. If we look at employment density – which might be a better measure of economic activity – coefficients are again close to the ones from our baseline results, especially when looking at NO_2. Between models (OLS or IV with our different instruments) the coefficients are relatively stable for almost all of the independent variables we look at.

An 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 Tab. B.20, 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.[28]

Finally, we control for the historical share of workers in industry and crafts, with results shown in Table B.21. This measure controls for the employment structure of an area in 1925. One concern with lagged population variables is that there may be unobserved factors driving both past and current population patterns. For instance, cities that grew in the past because of the presence of e.g. heavy industry may still have a large share of industrial plants with higher pollution than an economy based on services. Thus, pollution today and in the past may be higher in such areas compared to those with different employment structures. To mitigate this concern, we control for past employment structure in the IV regressions using the historic population density instrument.[29] In the uneven columns, we included the share of total employed individuals in industry and crafts, while even columns include only workers (i.e. exclude individuals working in administration or those who are self-employed, which is supposedly more common in crafts than in industry). Thus, the second measure is supposed to better capture the industrial employment in the region. Whatever control we include, the results are barely affected.

[28]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.

[29]It would be desirable to control for employment in differentiated industrial sectors, but we could not find these kind of variables over our sample of regions and time.

5.3 Threshold results

We now turn to the analysis of threshold violations. These have been the primary focus of recent policy debates, since German (and other European) cities and national governments have been sued for threshold violations.

For a first visual impression, Figures 5 and 6 show the number of transgressions of the daily mean threshold of PM_{10} ($50 \mu g/m^3$), and the NO_2 yearly mean ($20\ \mu g/m^3$)) by population density decile. The histograms suggest a clear positive association between density and threshold transgressions. In the right panel, we also show the same graph with total population, which shows a rather non-linear relationship: while there is no clear connection between city size and threshold violations, the twenty percent largest cities show more violations, especially when looking at NO_2 pollution.[30].

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 $\mu g/m^3$ 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 Tab. B.22).

Results for the transgressions of the 24-hour mean are shown in Table 7. The table shows that the probability of specific numbers of days with very high pollution readings 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 probability. In the case of PM_{10}, the probability is also significantly higher with the point estimate at 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 Tab. B.23, 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.

[30]For $PM_{2.5}$, corresponding figures do not show a clear pattern.

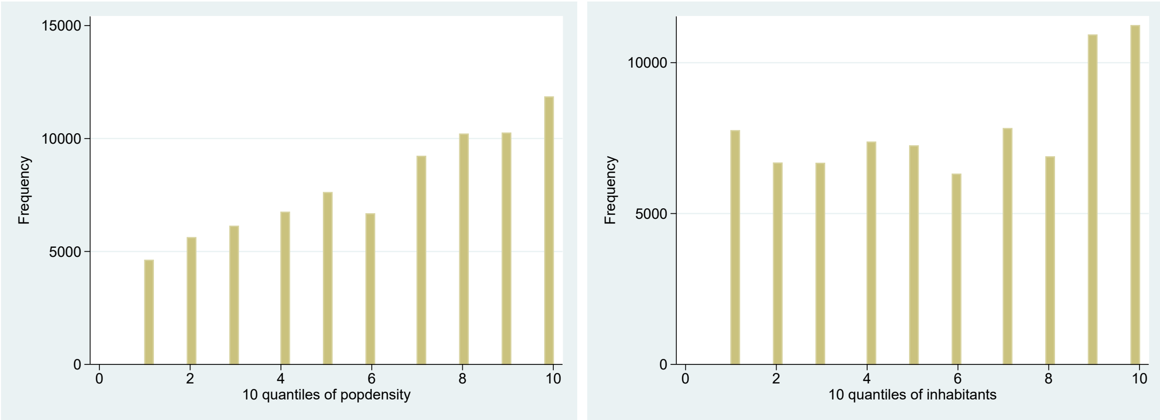


Figure 5: Histograms of PM_{10} daily mean threshold transgressions by deciles of population density and population

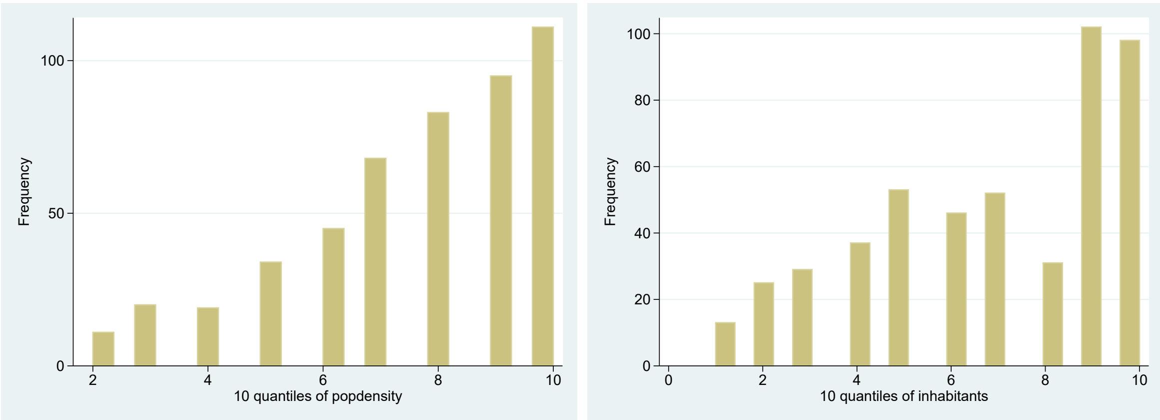


Figure 6: Histograms of NO_2 Yearly mean threshold transgressions by deciles of population density and population

Table 6: Probability of transgression of annual thresholds

	NO2			PM10			PM2.5		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	LPM	LPM Historical IV	LPM Soil IV	LPM	LPM Historical IV	LPM Soil IV	LPM	LPM Historical IV	LPM Soil IV
log(pop density)	0.148***	0.119***	0.235***	0.0984***	0.120***	0.0695*	-0.0112	-0.0242	-0.0131
	(0.0232)	(0.0334)	(0.0408)	(0.0216)	(0.0248)	(0.0367)	(0.0160)	(0.0173)	(0.0211)
distance to CBD	0.00143*	0.00102	0.00271***	0.00143	0.00145	0.000970	0.000296	-0.0000769	0.000265
	(0.000818)	(0.000932)	(0.000921)	(0.00102)	(0.00107)	(0.00110)	(0.000806)	(0.000807)	(0.000857)
suburban	-0.0376*	-0.0302*	-0.0412	0.221***	0.211***	0.222***	0.193***	0.142***	0.193***
	(0.0191)	(0.0179)	(0.0271)	(0.0403)	(0.0401)	(0.0408)	(0.0648)	(0.0517)	(0.0627)
urban	-0.0378	-0.0241	-0.0623*	0.278***	0.255***	0.287***	0.210***	0.163***	0.211***
	(0.0244)	(0.0265)	(0.0337)	(0.0466)	(0.0471)	(0.0476)	(0.0707)	(0.0597)	(0.0658)
industrial	-0.0000506	-0.00163	0.00383	0.264***	0.261***	0.263***	0.0426	0.0341*	0.0427*
	(0.0206)	(0.0186)	(0.0281)	(0.0422)	(0.0427)	(0.0421)	(0.0256)	(0.0195)	(0.0251)
traffic	0.549***	0.564***	0.544***	0.311***	0.313***	0.312***	0.00812	0.00629	0.00791
	(0.0427)	(0.0456)	(0.0436)	(0.0303)	(0.0320)	(0.0302)	(0.0139)	(0.0117)	(0.0134)
N	5663	5383	5663	4817	4565	4817	795	758	795
R^2	0.494	0.505	0.484	0.423	0.420	0.422	0.235	0.172	0.235
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.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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

	NO2			PM10			PM2.5		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	>17	>14	>9	>34	>29	>24	>34	>29	>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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusion

In this paper, we have used panel data for German districts to estimate the effect of population density on air pollution. To mitigate concerns about unobserved heterogeneity and omitted variables, we used both long difference regressions and instrumental variables. Our preferred estimates come from the IV regressions, where we instrumented population density with historical population and/or soil characteristics. We find that increasing population density by one standard deviation increases PM_{10} by about 3 percent and NO_2 by around 12 percent. The results for $PM_{2.5}$ are mostly insignificant due to the lower number of observations, while for O_3, OLS and IV results imply that denser cities are less prone to ozone pollution.

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.[31] Thus, our study makes some headway towards a more complete picture of agglomeration benefits and costs.

References

Ahlfeldt, G. and E. Pietrostefani (2019). The economic effects of density: A synthesis. CEPR Discussion Paper DP13440.

[31] See **?** 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.

Auffhammer, M., S. M. Hsiang, W. Schlenker, and A. Sobel (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy 7*(2), 181–198.

Auffhammer, M. and R. Kellogg (2011). Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review 101*(6), 2687–2722.

Bart, I. L. (2010). Urban sprawl and climate change: A statistical exploration of cause and effect, with policy options for the eu. *Land Use Policy 27*(2), 283–292.

Bauernschuster, S., T. Hener, and H. Rainer (2017, February). When labor disputes bring cities to a standstill: The impact of public transit strikes on traffic, accidents, air pollution, and health. *American Economic Journal: Economic Policy 9*(1), 1–37.

Blaudin de Thé, C., B. Carantino, and M. Lafourcade (2018). The carbon 'carprint' of suburbanization: New evidence from french cities. CEPR Discussion Papers 13086, CEPR.

Borck, R. (2017). Public transport and urban pollution. CESifo Working Paper Series 6606, CESifo.

Borck, R. and J. K. Brueckner (2018). Optimal energy taxation in cities. *Journal of the Association of Environmental and Resource Economists 5*(2), 481–516.

Borck, R. and M. Pflüger (2015). Green cities? Urbanization, trade and the environment. IZA DP No. 9104. Forthcoming, *Journal of Regional Science*.

Borck, R. and T. Tabuchi (2018). Pollution and city size: can cities be too small? *Journal of Economic Geography*, lby017. forthcoming.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Carozzi, F. and S. Roth (2019). Dirty density: Air quality and the density of American cities. Unpublished paper, LSE.

Ciccone, A. and R. E. Hall (1996, March). Productivity and the density of economic activity. *American Economic Review 86*(1), 54–70.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics*, pp. 15–66. University of Chicago Press.

Costa, S., J. Ferreira, C. Silveira, C. Costa, D. Lopes, H. Relvas, C. Borrego, P. Roebeling, A. I. Miranda, and J. Paulo Teixeira (2014). Integrating health on air quality assessment — Review report on health risks of two major european outdoor air pollutants: Pm and no2. *Journal of Toxicology and Environmental Health, Part B 17*(6), 307–340.

Ewing, R., R. Pendall, and D. Chen (2003). Measuring sprawl and its transportation impacts. *Transportation Research Record: Journal of the Transportation Research Board 1831*, 175–183.

Gehrsitz, M. (2017). The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management 83*, 121–144.

Glaeser, E. L. (1998). Are cities dying? *Journal of Economic Perspectives 12*(2), 139–160.

Glaeser, E. L. and M. E. Kahn (2010). The greenness of cities: carbon dioxide emissions and urban development. *Journal of Urban Economics 67*(3), 404–418.

Gudipudi, R., T. Fluschnik, A. G. C. Ros, C. Walther, and J. P. Kropp (2016). City density and CO2 efficiency. *Energy Policy 91*, 352–361.

Henderson, J. V. (1996). Effects of air quality regulation. *American Economic Review 86*(4), 789–813.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2018). The global distribution of economic activity: Nature, history, and the role of trade. *The Quarterly Journal of Economics 133*(1), 357–406.

Hilber, C. and C. Palmer (2014). Urban development and air pollution: Evidence from a global panel of cities. GRI Working Papers 175, Grantham Research Institute on Climate Change and the Environment.

Kahn, M. E. (2010). New evidence on trends in the cost of urban agglomeration. In E. L. Glaeser (Ed.), *Agglomeration economics*, pp. 339–354. University of Chicago Press.

Karathodorou, N., D. J. Graham, and R. B. Noland (2010). Estimating the effect of urban density on fuel demand. *Energy Economics 32*(1), 86–92.

Koh, H.-J., N. Riedel, and T. Böhm (2013). Do governments tax agglomeration rents? *Journal of Urban Economics 75*, 92–106.

Kosfeld, R. and A. Werner (2012). Deutsche Arbeitsmarktregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011. *Raumforschung und Raumordnung 70*(1), 49–64.

Lamsal, L., R. Martin, D. D. Parrish, and N. A. Krotkov (2013). Scaling relationship for NO2 pollution and urban population size: a satellite perspective. *Environmental Science & Technology 47*(14), 7855–7861.

Larson, W. and A. Yezer (2015). The energy implications of city size and density. *Journal of Urban Economics 90*, 35–49.

Newman, P. W. and J. R. Kenworthy (1989). Gasoline consumption and cities: A comparison of US cities with a global survey. *Journal of the American Planning Association 55*(1), 24–37.

Oliveira, E. A., J. S. Andrade Jr, and H. A. Makse (2014). Large cities are less green. *Scientific Reports 4*, Article number: 4235.

Panagos, P., M. Van Liedekerke, A. Jones, and L. Montanarella (2012). European soil data centre: Response to european policy support and public data requirements. *Land use policy 29*(2), 329–338.

Pope III, C. A. and D. W. Dockery (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air & Waste Management Association 56*(6), 709–742.

Redding, S. J. and D. M. Sturm (2008). The costs of remoteness: Evidence from german division and reunification. *American Economic Review 98*(5), 1766–1797.

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

Stone, B. (2008). Urban sprawl and air quality in large US cities. *Journal of Environmental Management 86*(4), 688–698.

WHO, O. (2015). Economic cost of the health impact of air pollution in Europe: Clean air, health and wealth. *WHO Regional Office for Europe. Dostępny pod adresem: http://www. euro. who. int/en/media-centre/events/events/2015/04/ehp-mid-termreview/publications/economic-cost-of-the-health-impact-of-air-pollution-in.*

Wolff, H. (2014). Keep your clunker in the suburb: Low-emission zones and adoption of green vehicles. *Economic Journal 124*(578), F481–F512.

World Health Organization (2003). Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide: Report on a WHO working group, bonn, germany 13-15 january 2003.

World Health Organization, UNAIDS, et al. (2006). *Air Quality Guidelines: Global Update 2005.* World Health Organization.

Zhou, Y., J. I. Levy, J. S. Evans, and J. K. Hammitt (2006). The influence of geographic location on population exposure to emissions from power plants throughout China. *Environment International 32*(3), 365–373.

Appendix

A A model

Consider a circular monocentric city where N residents commute to the CBD for work. A resident household living at x km from the CBD incurs round-trip commuting costs tx. Household utility is $v(c,q) = c^{1-\alpha} q^{\alpha} E^{-\beta}$, where c is non-housing consumption, q consumption of housing floor space in square meters, and E 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 $q = \alpha u^{\frac{1}{\alpha}} e^{\frac{\beta}{\alpha}} (y - tx)^{1-\frac{1}{\alpha}}$, and the bid rent, i.e. the maximum willingness to pay per unit of housing floor space, $p = u^{-1/\alpha}(y - tx)^{\frac{1}{\alpha}} E^{-\frac{\beta}{\alpha}}$.

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 $h = S^{\theta}$, where $S = K/L$ 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, $S = \theta^{\frac{1}{1-\theta}} u^{\frac{1}{\alpha(\theta-1)}} (y - tx)^{\frac{1}{\alpha - \alpha\theta}} E^{\frac{\beta}{\alpha(\theta-1)}}$, and the land rent function at distance x, $R = \left(\theta^{\frac{\theta}{1-\theta}} - \theta^{\frac{1}{1-\theta}}\right) u^{\frac{1}{\alpha(\theta-1)}} (y - tx)^{\frac{1}{\alpha-\alpha\theta}} E^{\frac{\beta}{\alpha(\theta-1)}}$.

The equilibrium in the closed city is defined by the two equations

$$R(\bar{x}, u, E) = R_A \tag{A.1}$$
$$\int_0^{\bar{x}} \frac{h(x,u,E)}{q(x,u,E)} dx = N, \tag{A.2}$$

where $\bar{x}$ is the distance from the city border to the CBD and R_A is the agricultural land rent.

Solving (A.1) and (A.2) gives the residents utility level $u()$ and the city border, $\bar{x}()$, both of which can not be solved analytically. Pollution is composed of pollution from commuting and residential energy use, weighted by the respective emissions factors.

Letting the emissions intensities of commuting and housing be e_C and e_H, total emissions are

$$E = e_C C + e_H H \tag{A.3}$$
$$C = \int_0^{\bar{x}} x D(x) 2\pi x dx \tag{A.4}$$
$$H = \int_0^{\bar{x}} h(x) 2\pi x dx. \tag{A.5}$$

Finally, assume for simplicity that the concentration of air pollution is given by total emissions divided by land area.[32] Then concentration is given by $\mathcal{C} = E/(\pi\bar{x}^2)$.

How then does concentration change with population density? There are two ways of inducing an increase in density in the model. First, we could increase population, which would increase the city border and lead to an increase of density over the entire city area. Consequently, total pollution from transport rises, as residents face longer average commutes. While residents live in smaller dwellings due to the increased pressure on the housing market, total residential energy use in the simulation rises. Finally, total land area rises, but in our example, concentration rises as pollution increases faster than land area.

Secondly, suppose the agricultural land rent R_A rises. This reduces the city border and increases density in the entire city. Now, pollution falls as residents face shorter average commutes and average dwelling size falls, again due to increased competition for central land. However, pollution falls less than land area in the simulation, so concentration rises in our example again.

[32] 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.

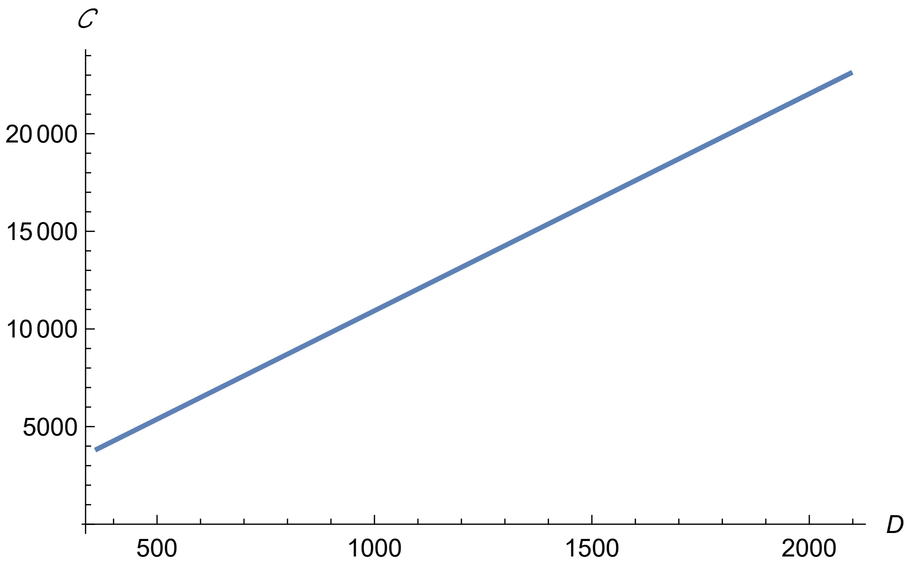


Figure A.1: Effect of population increase on pollutant concentration

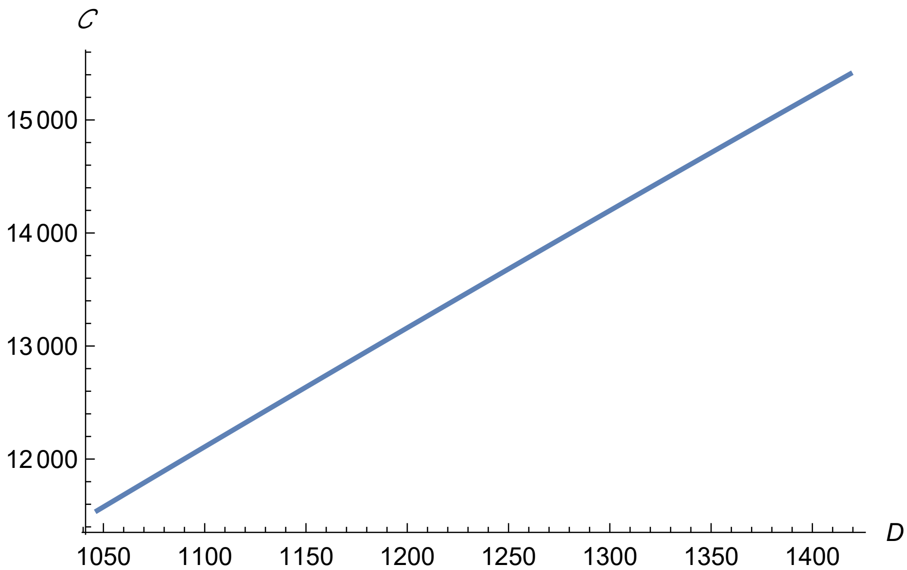


Figure A.2: Effect of land rent increase on pollutant concentration

B 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 kilometers and a maximum station altitude difference of 200 meters.[33] 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

[33] 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.

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).[34] The data comes in raster format of 1km$\times$1km 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 to 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 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.

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

C Further results and robustness checks

Table B.1: First stage regressions for the historical population density 1910, the soil IV, and the IV including density in 1910 and soil (NO_2 and PM_{10})

	NO_2			PM_{10}		
	(1) Density 1910	(2) Soil	(3) Density 1910 + Soil	(4) Density 1910	(5) Soil	(6) Density 1910 + Soil
logdensity1910	0.241***		0.245***	0.254***		0.258***
	(0.0136)		(0.0195)	(0.0130)		(0.0182)
N	5301	5575	5301	4407	4648	4407
R^2	0.744	0.749	0.749	0.459	0.468	0.463
Districts	269	269	269	247	247	247
Soil Characteristics	No	Yes	Yes	No	Yes	Yes
First-stage Statistic	312.8	20.62	57.95	385.4	13.96	110.4
Overidentification		30.52	32.83		37.39	37.36
Hansen p-stat		0.439	0.377		0.136	0.167

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 B.2: First stage regressions for the historical population density 1910, the soil IV, and the IV including density in 1910 and soil ($PM_{2.5}$ and O_3)

	$PM_{2.5}$			O_3		
	(1) Density 1910	(2) Soil	(3) Density 1910 + Soil	(4) Density 1910	(5) Soil	(6) Density 1910 + Soil
logdensity1910	0.248***		0.296***	0.257***		0.265***
	(0.0242)		(0.0322)	(0.0133)		(0.0197)
N	758	795	758	3588	3776	3588
R^2	0.223	0.246	0.228	0.416	0.436	0.431
Districts	109	109	109	251	251	251
Soil Characteristics	No	Yes	Yes	No	Yes	Yes
First-stage Statistic	105.3	59.11	176.6	374.1	42.63	81.32
Hansen p-stat		0.135	.		0.0599	0.0918

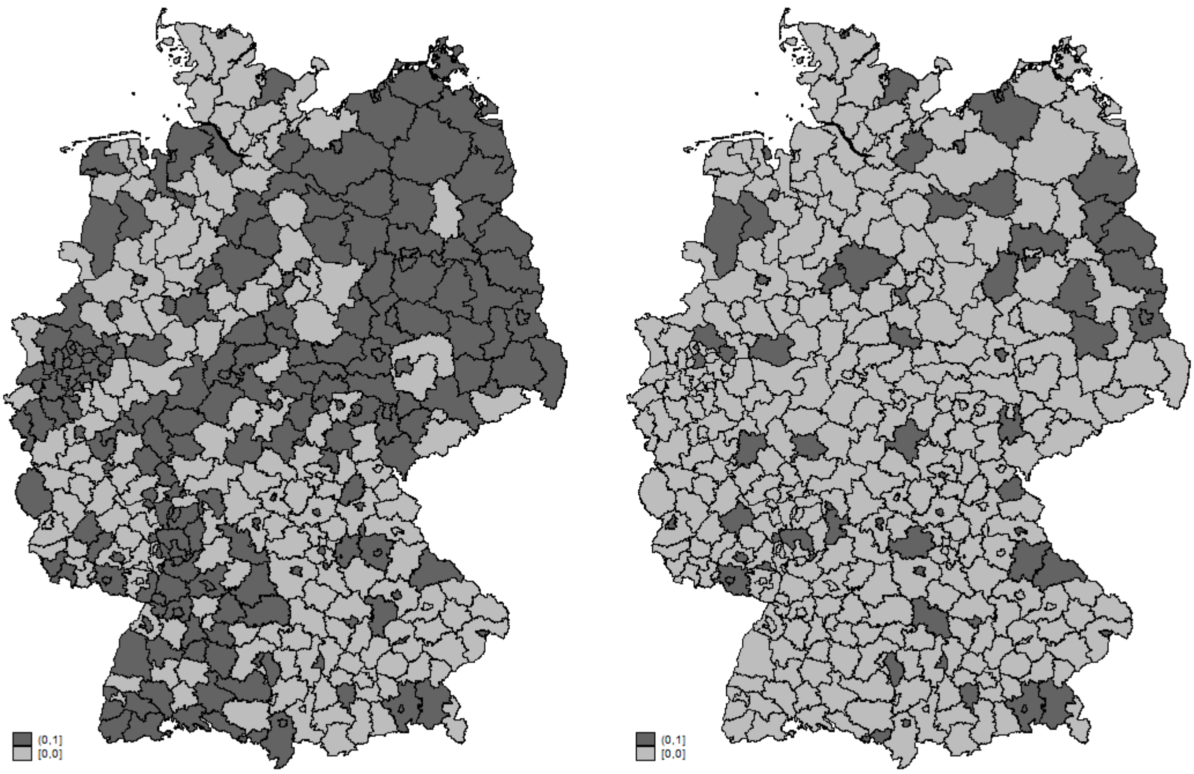
Standard errors in parantheses are clustered at labor market region - year (OLS) and labor market region (IV) level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.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
R^2	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

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



(a) Sample of PM_{10} (b) Sample of $PM_{2.5}$

Figure B.1: Districts included into PM_{10} and $PM_{2.5}$ regression analyses

Table B.4: OLS regression for NO_2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pop density)	0.588***	0.591***	0.281***	0.281***	0.239***	0.229***	0.266***	0.215***
	(0.0176)	(0.0176)	(0.0137)	(0.0141)	(0.0144)	(0.0126)	(0.0146)	(0.0164)
Distance to CBD			-0.00384***	-0.00390***	-0.00415***	-0.00354***	-0.00195***	-0.00236***
			(0.000466)	(0.000465)	(0.000449)	(0.000430)	(0.000434)	(0.000427)
Suburban			0.345***	0.343***	0.345***	0.354***	0.279***	0.335***
			(0.0152)	(0.0151)	(0.0151)	(0.0144)	(0.0179)	(0.0178)
Urban			0.513***	0.512***	0.524***	0.573***	0.471***	0.489***
			(0.0185)	(0.0185)	(0.0182)	(0.0167)	(0.0210)	(0.0201)
Industrial			0.0888***	0.0853***	0.0750***	0.115***	0.106***	0.110***
			(0.0133)	(0.0143)	(0.0132)	(0.0135)	(0.0150)	(0.0145)
Traffic			0.661***	0.660***	0.660***	0.651***	0.615***	0.633***
			(0.0126)	(0.0127)	(0.0126)	(0.0121)	(0.0115)	(0.0119)
Steinkohle				-0.00413				
				(0.0144)				
Braunkohle				0.0505**				
				(0.0212)				
Distance to coal plant					-0.00152***			
					(0.000183)			
Av. GDP						-0.0110	0.0403***	0.0292
						(0.0154)	(0.0154)	(0.0180)
Av. Income						0.401***	0.395***	0.243***
						(0.0508)	(0.0546)	(0.0564)
Unemployment share						-1.153***	-1.372***	-0.611**
						(0.192)	(0.211)	(0.240)
Green Voters							-1.081***	-0.873***
							(0.155)	(0.155)
Env. Zone							0.0461***	0.0557***
							(0.00760)	(0.00792)
Distance to Street							-0.118***	-0.100***
							(0.0112)	(0.0109)
N	5575	5575	5575	5575	5575	5489	4905	4905
R^2	0.303	0.305	0.749	0.749	0.754	0.767	0.771	0.795
Districts	269	269	269	269	269	269	269	269
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	Yes	No	No	No	No
Distance to coal plant	No	No	No	No	Yes	No	No	No

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 B.5: OLS regression for PM_{10}

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pop density)	0.163***	0.164***	0.0747***	0.0642***	0.0766***	0.120***	0.133***	0.0774***
	(0.00637)	(0.00637)	(0.00691)	(0.00763)	(0.00760)	(0.00809)	(0.00996)	(0.00986)
Distance to CBD			0.000692**	0.000542*	0.000706**	0.000320	0.000408	-0.000665**
			(0.000284)	(0.000290)	(0.000287)	(0.000268)	(0.000307)	(0.000280)
Suburban			0.103***	0.102***	0.103***	0.118***	0.0957***	0.0944***
			(0.00882)	(0.00882)	(0.00882)	(0.00873)	(0.0105)	(0.0102)
Urban			0.168***	0.167***	0.168***	0.171***	0.148***	0.166***
			(0.01000)	(0.0101)	(0.0101)	(0.00987)	(0.0128)	(0.0117)
Industrial			0.136***	0.130***	0.136***	0.124***	0.126***	0.130***
			(0.0126)	(0.0127)	(0.0125)	(0.0121)	(0.0137)	(0.0114)
Traffic			0.260***	0.262***	0.260***	0.254***	0.250***	0.250***
			(0.00670)	(0.00681)	(0.00671)	(0.00659)	(0.00710)	(0.00656)
Steinkohle				0.0274***				
				(0.00910)				
Braunkohle				0.0202				
				(0.0209)				
Distance to coal plant					0.0000731			
					(0.000141)			
Av. GDP						-0.117***	-0.0861***	-0.0395***
						(0.0104)	(0.0113)	(0.0115)
Av. Income						0.125***	0.155***	0.0458
						(0.0361)	(0.0410)	(0.0400)
Unemployment share						0.761***	0.538***	-0.274*
						(0.123)	(0.143)	(0.150)
Green Voters							-0.959***	-0.961***
							(0.128)	(0.125)
Env. Zone							0.00602	0.00361
							(0.00547)	(0.00479)
Distance to Street							-0.0359***	-0.0157**
							(0.00721)	(0.00615)
N	4648	4648	4648	4648	4648	4570	4137	4137
R^2	0.142	0.143	0.469	0.470	0.469	0.498	0.491	0.587
Districts	247	247	247	247	247	247	247	247
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	Yes	No	No	No	No
Distance to coal plant	No	No	No	No	Yes	No	No	No

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 B.6: OLS regression for $PM_{2.5}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pop density)	0.0898***	0.0912***	0.0315**	0.00289	0.0298*	0.0550***	0.0665***	0.00197
	(0.0109)	(0.0110)	(0.0160)	(0.0157)	(0.0176)	(0.0207)	(0.0204)	(0.0174)
Distance to CBD			0.00105	0.000617	0.00104	0.000810	0.00103	-0.000435
			(0.000641)	(0.000628)	(0.000659)	(0.000609)	(0.000642)	(0.000498)
Suburban			0.168***	0.167***	0.167***	0.176***	0.116***	0.0935***
			(0.0226)	(0.0226)	(0.0230)	(0.0218)	(0.0283)	(0.0205)
Urban			0.209***	0.207***	0.209***	0.200***	0.138***	0.135***
			(0.0244)	(0.0245)	(0.0245)	(0.0244)	(0.0319)	(0.0244)
Industrial			0.0699***	0.0534***	0.0681***	0.0717***	0.0363	0.0526***
			(0.0202)	(0.0200)	(0.0207)	(0.0195)	(0.0251)	(0.0197)
Traffic			0.115***	0.120***	0.115***	0.128***	0.120***	0.136***
			(0.0162)	(0.0165)	(0.0160)	(0.0166)	(0.0190)	(0.0164)
Steinkohle				0.0969***				
				(0.0170)				
Braunkohle				0.188***				
				(0.0265)				
Distance to coal plant					-0.0000853			
					(0.000292)			
Av. GDP						-0.0526**	-0.0492**	-0.0143
						(0.0245)	(0.0242)	(0.0196)
Av. Income						0.0601	-0.0652	-0.464***
						(0.0864)	(0.0857)	(0.0800)
Unemployment share						1.222***	0.422	-1.594***
						(0.376)	(0.380)	(0.367)
Green Voters							-0.960***	-0.00407
							(0.281)	(0.222)
Env. Zone							0.0275***	0.0313***
							(0.00853)	(0.00654)
Distance to Street							-0.0478***	-0.00819
							(0.0167)	(0.0114)
N	795	795	795	795	795	773	719	719
R^2	0.068	0.069	0.246	0.278	0.246	0.280	0.307	0.600
Districts	109	109	109	109	109	109	109	109
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	Yes	No	No	No	No
Distance to coal plant	No	No	No	No	Yes	No	No	No

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 B.7: OLS regression for O_3

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pop density)	-0.269***	-0.271***	-0.176***	-0.159***	-0.141***	-0.131***	-0.149***	-0.0722***
	(0.00994)	(0.00998)	(0.00938)	(0.00988)	(0.00918)	(0.00909)	(0.0121)	(0.0136)
Distance to CBD			0.00274***	0.00314***	0.00274***	0.00214***	0.00155***	0.00137***
			(0.000297)	(0.000306)	(0.000291)	(0.000261)	(0.000279)	(0.000293)
Suburban			-0.167***	-0.159***	-0.168***	-0.175***	-0.159***	-0.187***
			(0.0111)	(0.0111)	(0.0111)	(0.0110)	(0.0124)	(0.0138)
Urban			-0.225***	-0.220***	-0.235***	-0.276***	-0.262***	-0.278***
			(0.0119)	(0.0121)	(0.0118)	(0.0114)	(0.0146)	(0.0146)
Industrial			-0.0688***	-0.0505***	-0.0635***	-0.111***	-0.110***	-0.0688***
			(0.0115)	(0.0122)	(0.0108)	(0.0121)	(0.0129)	(0.0140)
Traffic			-0.233***	-0.236***	-0.231***	-0.220***	-0.222***	-0.266***
			(0.0162)	(0.0163)	(0.0155)	(0.0140)	(0.0148)	(0.0148)
Steinkohle				-0.0342***				
				(0.0127)				
Braunkohle				-0.169***				
				(0.0302)				
Distance to coal plant					0.00133***			
					(0.000150)			
Av. GDP						-0.00445	-0.0285*	-0.0677***
						(0.0138)	(0.0146)	(0.0160)
Av. Income						-0.243***	-0.246***	-0.148**
						(0.0600)	(0.0611)	(0.0581)
Unemployment share						1.450***	1.839***	1.173***
						(0.158)	(0.174)	(0.185)
Green Voters							0.770***	0.673***
							(0.147)	(0.161)
Env. Zone							0.00449	0.00618
							(0.00719)	(0.00673)
Distance to Street							0.0530***	0.0355***
							(0.00763)	(0.00791)
N	3776	3776	3776	3776	3776	3722	3438	3438
R^2	0.258	0.260	0.437	0.443	0.452	0.499	0.511	0.562
Districts	251	251	251	251	251	251	251	251
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	Yes	No	No	No	No
Distance to coal plant	No	No	No	No	Yes	No	No	No

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 B.8: Different specifications for NO_2 regressions

	(1) OLS	(2) IV	(3) IV Soil	(4) OLS	(5) IV	(6) IV Soil	(7) OLS	(8) IV	(9) IV Soil
log(pop density)	0.588***	0.570***	0.520***	0.239***	0.129**	0.214***	0.266***	0.226***	0.302***
	(0.0176)	(0.0889)	(0.0662)	(0.0144)	(0.0612)	(0.0615)	(0.0146)	(0.0820)	(0.0916)
Distance to CBD				-0.00415***	-0.00590***	-0.00450***	-0.00195***	-0.00233*	-0.00175
				(0.000449)	(0.00155)	(0.00141)	(0.000434)	(0.00136)	(0.00127)
Suburban				0.345***	0.340***	0.347***	0.279***	0.271***	0.279***
				(0.0151)	(0.0461)	(0.0443)	(0.0179)	(0.0516)	(0.0498)
Urban				0.524***	0.554***	0.533***	0.471***	0.474***	0.466***
				(0.0182)	(0.0578)	(0.0562)	(0.0210)	(0.0640)	(0.0626)
Industrial				0.0750***	0.0700*	0.0737**	0.106***	0.106***	0.105***
				(0.0132)	(0.0387)	(0.0357)	(0.0150)	(0.0407)	(0.0377)
Traffic				0.660***	0.668***	0.661***	0.615***	0.620***	0.615***
				(0.0126)	(0.0418)	(0.0395)	(0.0115)	(0.0371)	(0.0356)
Distance to coal plant				-0.00152***	-0.00217***	-0.00169***			
				(0.000183)	(0.000678)	(0.000646)			
Av. GDP							0.0403***	0.0699	0.0225
							(0.0154)	(0.0632)	(0.0590)
Av. Income							0.395***	0.352**	0.373**
							(0.0546)	(0.138)	(0.157)
Unemployment share							-1.372***	-1.408***	-1.503**
							(0.211)	(0.496)	(0.614)
Green Voters							-1.081***	-0.940	-1.230**
							(0.155)	(0.581)	(0.573)
Env. Zone							0.0461***	0.0482***	0.0430***
							(0.00760)	(0.0139)	(0.0140)
Distance to Street							-0.118***	-0.116***	-0.115***
							(0.0112)	(0.0369)	(0.0358)
N	5575	5301	5575	5575	5301	5575	4905	4631	4905
R^2	0.303	0.305	0.299	0.754	0.748	0.753	0.771	0.771	0.771
Districts	269	269	269	269	269	269	269	269	269
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	No	No	No	No	No	No
Distance to coal plant	No	No	No	Yes	Yes	Yes	No	No	No

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 B.9: Different specifications for PM_{10} regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	IV	IV Soil	OLS	IV	IV Soil	OLS	IV	IV Soil
log(pop density)	0.163***	0.208***	0.127***	0.0766***	0.111***	0.0626	0.133***	0.166***	0.109*
	(0.00637)	(0.0201)	(0.0344)	(0.00760)	(0.0298)	(0.0441)	(0.00996)	(0.0464)	(0.0578)
Distance to CBD				0.000706**	0.000929	0.000509	0.000408	0.000269	0.000267
				(0.000287)	(0.000861)	(0.000925)	(0.000307)	(0.000861)	(0.000881)
Suburban				0.103***	0.0921***	0.104***	0.0957***	0.0903***	0.0952***
				(0.00882)	(0.0232)	(0.0245)	(0.0105)	(0.0278)	(0.0286)
Urban				0.168***	0.140***	0.172***	0.148***	0.136***	0.151***
				(0.0101)	(0.0298)	(0.0314)	(0.0128)	(0.0380)	(0.0369)
Industrial				0.136***	0.134***	0.135***	0.126***	0.121***	0.127***
				(0.0125)	(0.0354)	(0.0361)	(0.0137)	(0.0382)	(0.0376)
Traffic				0.260***	0.261***	0.260***	0.250***	0.249***	0.250***
				(0.00671)	(0.0184)	(0.0177)	(0.00710)	(0.0195)	(0.0189)
Distance to coal plant				0.0000731	0.000365	-0.0000225			
				(0.000141)	(0.000464)	(0.000504)			
Av. GDP							-0.0861***	-0.100**	-0.0740**
							(0.0113)	(0.0474)	(0.0377)
Av. Income							0.155***	0.131	0.169
							(0.0410)	(0.104)	(0.111)
Unemployment share							0.538***	0.392	0.629
							(0.143)	(0.342)	(0.431)
Green Voters							-0.959***	-1.164***	-0.851**
							(0.128)	(0.376)	(0.418)
Env. Zone							0.00602	0.00201	0.00847
							(0.00547)	(0.00915)	(0.00996)
Distance to Street							-0.0359***	-0.0261	-0.0374*
							(0.00721)	(0.0217)	(0.0226)
N	4648	4407	4648	4648	4407	4648	4137	3896	4137
R^2	0.142	0.131	0.135	0.469	0.459	0.468	0.491	0.479	0.490
Districts	247	247	247	247	247	247	247	247	247
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	No	No	No	No	No	No
Distance to coal plant	No	No	No	Yes	Yes	Yes	No	No	No

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 B.10: Different specifications for $PM_{2.5}$ regressions

	(1) OLS	(2) IV	(3) IV Soil	(4) OLS	(5) IV	(6) IV Soil	(7) OLS	(8) IV	(9) IV Soil
log(pop density)	0.0898***	0.108***	0.0759**	0.0298*	0.0739	0.0447	0.0665***	0.0582	0.0706
	(0.0109)	(0.0311)	(0.0385)	(0.0176)	(0.0620)	(0.0631)	(0.0204)	(0.0671)	(0.0859)
Distance to CBD				0.00104	0.00135	0.00126	0.00103	0.000809	0.00105
				(0.000659)	(0.00156)	(0.00156)	(0.000642)	(0.00153)	(0.00147)
Suburban				0.167***	0.162***	0.166***	0.116***	0.116**	0.116**
				(0.0230)	(0.0528)	(0.0483)	(0.0283)	(0.0577)	(0.0576)
Urban				0.209***	0.177***	0.200***	0.138***	0.135**	0.137**
				(0.0245)	(0.0558)	(0.0560)	(0.0319)	(0.0653)	(0.0688)
Industrial				0.0681***	0.0781*	0.0690	0.0363	0.0368	0.0365
				(0.0207)	(0.0434)	(0.0420)	(0.0251)	(0.0489)	(0.0491)
Traffic				0.115***	0.115***	0.117***	0.120***	0.120***	0.121***
				(0.0160)	(0.0378)	(0.0371)	(0.0190)	(0.0449)	(0.0459)
Distance to coal plant				-0.0000853	0.000162	-0.0000167			
				(0.000292)	(0.000740)	(0.000725)			
Av. GDP							-0.0492**	-0.0380	-0.0516
							(0.0242)	(0.0631)	(0.0610)
Av. Income							-0.0652	-0.0520	-0.0663
							(0.0857)	(0.172)	(0.167)
Unemployment share							0.422	0.540	0.404
							(0.380)	(0.872)	(0.785)
Green Voters							-0.960***	-0.922	-0.979
							(0.281)	(0.684)	(0.774)
Env. Zone							0.0275***	0.0279*	0.0271
							(0.00853)	(0.0153)	(0.0170)
Distance to Street							-0.0478***	-0.0454	-0.0479
							(0.0167)	(0.0368)	(0.0352)
N	795	758	795	795	758	795	719	682	719
R^2	0.068	0.066	0.066	0.246	0.223	0.245	0.307	0.289	0.307
Districts	109	109	109	109	109	109	109	109	109
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	No	No	No	No	No	No
Distance to coal plant	No	No	No	Yes	Yes	Yes	No	No	No

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 B.11: Different specifications for O_3 regressions

	(1) OLS	(2) IV	(3) IV Soil	(4) OLS	(5) IV	(6) IV Soil	(7) OLS	(8) IV	(9) IV Soil
log(pop density)	-0.269***	-0.217***	-0.282***	-0.141***	-0.0463	-0.130***	-0.149***	-0.0639	-0.213***
	(0.00994)	(0.0427)	(0.0393)	(0.00918)	(0.0326)	(0.0438)	(0.0121)	(0.0621)	(0.0819)
Distance to CBD				0.00274***	0.00391***	0.00286***	0.00155***	0.00187**	0.00127
				(0.000291)	(0.000972)	(0.000966)	(0.000279)	(0.000810)	(0.000872)
Suburban				-0.168***	-0.167***	-0.169***	-0.159***	-0.150***	-0.159***
				(0.0111)	(0.0347)	(0.0340)	(0.0124)	(0.0375)	(0.0367)
Urban				-0.235***	-0.258***	-0.238***	-0.262***	-0.267***	-0.251***
				(0.0118)	(0.0379)	(0.0390)	(0.0146)	(0.0442)	(0.0489)
Industrial				-0.0635***	-0.0402	-0.0616**	-0.110***	-0.0982***	-0.111***
				(0.0108)	(0.0298)	(0.0276)	(0.0129)	(0.0331)	(0.0319)
Traffic				-0.231***	-0.235***	-0.231***	-0.222***	-0.231***	-0.221***
				(0.0155)	(0.0373)	(0.0346)	(0.0148)	(0.0317)	(0.0322)
Distance to coal plant				0.00133***	0.00189***	0.00141***			
				(0.000150)	(0.000520)	(0.000506)			
Av. GDP							-0.0285*	-0.102*	0.00502
							(0.0146)	(0.0529)	(0.0567)
Av. Income							-0.246***	-0.304	-0.199
							(0.0611)	(0.200)	(0.197)
Unemployment share							1.839***	1.529***	2.066***
							(0.174)	(0.504)	(0.578)
Green Voters							0.770***	0.554	1.036*
							(0.147)	(0.524)	(0.546)
Env. Zone							0.00449	-0.00386	0.0117
							(0.00719)	(0.0143)	(0.0144)
Distance to Street							0.0530***	0.0540**	0.0506**
							(0.00763)	(0.0253)	(0.0250)
N	3776	3588	3776	3776	3588	3776	3438	3250	3438
R^2	0.258	0.241	0.257	0.452	0.431	0.452	0.511	0.510	0.506
Districts	251	251	251	251	251	251	251	251	251
Weather	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Coal plant in district	No	No	No	No	No	No	No	No	No
Distance to coal plant	No	No	No	Yes	Yes	Yes	No	No	No

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 B.12: Regression results for district-free cities

	NO_2			PM_{10}		
	(1) OLS	(2) IV Historical	(3) IV Soil	(4) OLS	(5) IV Historical	(6) IV Soil
log(pop density)	0.537***	0.419**	0.553***	0.181***	0.419***	0.217***
	(0.0286)	(0.184)	(0.108)	(0.0142)	(0.140)	(0.0648)
Distance to CBD	-0.0281***	-0.0240**	-0.0288***	-0.00230**	-0.0122**	-0.00391
	(0.00245)	(0.0100)	(0.00795)	(0.00104)	(0.00546)	(0.00335)
Suburban	0.365***	0.354***	0.363***	0.0839***	0.0457	0.0770**
	(0.0358)	(0.0982)	(0.0997)	(0.0147)	(0.0560)	(0.0300)
Urban	0.504***	0.538***	0.501***	0.125***	0.0422	0.112***
	(0.0328)	(0.0969)	(0.0960)	(0.0132)	(0.0735)	(0.0268)
Industrial	0.279***	0.265***	0.280***	0.176***	0.194***	0.181***
	(0.0223)	(0.0662)	(0.0700)	(0.0200)	(0.0712)	(0.0639)
Traffic	0.614***	0.611***	0.614***	0.274***	0.272***	0.273***
	(0.0137)	(0.0410)	(0.0399)	(0.00752)	(0.0212)	(0.0203)
N	2661	2506	2661	2305	2148	2305
R^2	0.700	0.705	0.700	0.466	0.390	0.464
Districts	88	88	88	85	85	85
Weather	Yes	Yes	Yes	Yes	Yes	Yes

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 B.13: Regression results for district-free cities

	$PM_{2.5}$			O_3		
	(1) OLS	(2) IV Historical	(3) IV Soil	(4) OLS	(5) IV Historical	(6) IV Soil
log(pop density)	0.0625**	0.246	0.0645	-0.309***	0.0724	-0.372***
	(0.0247)	(0.170)	(0.0707)	(0.0232)	(0.181)	(0.0910)
Distance to CBD	0.00865**	0.00491	0.00861	0.00488***	-0.0103	0.00713
	(0.00393)	(0.00635)	(0.00706)	(0.00187)	(0.00915)	(0.00530)
Suburban	0.294***	0.171	0.293***	-0.191***	-0.181**	-0.191***
	(0.0470)	(0.124)	(0.0770)	(0.0214)	(0.0727)	(0.0634)
Urban	0.358***	0.216	0.357***	-0.259***	-0.318***	-0.252***
	(0.0460)	(0.132)	(0.0775)	(0.0212)	(0.0853)	(0.0636)
Industrial	0.0896***	0.0832**	0.0896***	-0.268***	-0.152***	-0.282***
	(0.0277)	(0.0410)	(0.0340)	(0.0138)	(0.0534)	(0.0389)
Traffic	0.131***	0.131***	0.131***	-0.253***	-0.271***	-0.252***
	(0.0176)	(0.0348)	(0.0358)	(0.0202)	(0.0582)	(0.0503)
N	433	430	433	1301	1194	1301
R^2	0.262	0.185	0.262	0.325	0.195	0.320
Districts	51	51	51	72	72	72
Weather	Yes	Yes	Yes	Yes	Yes	Yes

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 B.14: Labor Market Regions Regressions (NO_2 and PM_{10})

	NO_2			PM_{10}		
	(1) OLS	(2) IV Historic	(3) IV Soil	(4) OLS	(5) IV Historic	(6) IV Soil
log(pop density)	0.315***	0.231***	0.300***	0.0966***	0.125***	0.0765
	(0.0656)	(0.0729)	(0.0864)	(0.0246)	(0.0220)	(0.0511)
Distance to CBD	-0.00729***	-0.00755***	-0.00733***	-0.0000962	0.0000146	-0.000174
	(0.00150)	(0.00147)	(0.00151)	(0.000707)	(0.000713)	(0.000702)
Urban	0.565***	0.576***	0.567***	0.181***	0.176***	0.184***
	(0.0501)	(0.0502)	(0.0515)	(0.0268)	(0.0264)	(0.0269)
Traffic	0.670***	0.671***	0.670***	0.265***	0.266***	0.265***
	(0.0365)	(0.0372)	(0.0363)	(0.0168)	(0.0166)	(0.0169)
N	5575	5575	5575	4648	4648	4648
R^2	0.748	0.744	0.748	0.476	0.473	0.475
Labor Market Regions	128	128	128	125	125	125
Weather	Yes	Yes	Yes	Yes	Yes	Yes

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 B.15: Labor Market Regions Regressions ($PM_{2.5}$ and O_3)

	$PM_{2.5}$			O_3		
	(1) OLS	(2) IV Historic	(3) IV Soil	(4) OLS	(5) IV Historic	(6) IV Soil
log(pop density)	0.0731*	0.110***	0.0684	-0.233***	-0.190***	-0.325***
	(0.0399)	(0.0408)	(0.0704)	(0.0451)	(0.0561)	(0.0676)
Distance to CBD	0.000749	0.000844	0.000736	0.00425***	0.00442***	0.00389***
	(0.00119)	(0.00116)	(0.00118)	(0.00101)	(0.00101)	(0.00109)
Urban	0.206***	0.193***	0.207***	-0.259***	-0.263***	-0.251***
	(0.0441)	(0.0439)	(0.0444)	(0.0359)	(0.0354)	(0.0372)
Traffic	0.116***	0.118***	0.115***	-0.253***	-0.249***	-0.263***
	(0.0378)	(0.0368)	(0.0367)	(0.0349)	(0.0354)	(0.0351)
N	795	795	795	3776	3776	3776
R^2	0.263	0.258	0.263	0.463	0.459	0.446
Labor Market Regions	77	77	77	126	126	126
Weather	Yes	Yes	Yes	Yes	Yes	Yes

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 B.16: Alternative independent variables for NO_2

	Alternative population density			Log of population			Log of Employed per area		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil
logaltdensity	0.112***	0.0754***	0.104***						
	(0.00577)	(0.0203)	(0.0234)						
logpopulation				0.0961***	0.117**	0.0773**			
				(0.00752)	(0.0518)	(0.0308)			
logempldensity							0.211***	0.149***	0.244***
							(0.00890)	(0.0374)	(0.0417)
Distance to CBD	-0.00430***	-0.00579***	-0.00456***	-0.00898***	-0.00915***	-0.00884***	-0.00333***	-0.00500***	-0.00256*
	(0.000478)	(0.00158)	(0.00150)	(0.000451)	(0.00152)	(0.00142)	(0.000476)	(0.00163)	(0.00146)
Suburban	0.343***	0.339***	0.345***	0.360***	0.346***	0.362***	0.338***	0.335***	0.332***
	(0.0154)	(0.0477)	(0.0459)	(0.0155)	(0.0483)	(0.0464)	(0.0151)	(0.0459)	(0.0452)
Urban	0.518***	0.541***	0.524***	0.573***	0.557***	0.580***	0.508***	0.533***	0.492***
	(0.0185)	(0.0586)	(0.0579)	(0.0175)	(0.0627)	(0.0534)	(0.0183)	(0.0577)	(0.0594)
Industrial	0.0630***	0.0714	0.0650	0.0666***	0.0562	0.0717	0.0979***	0.0926**	0.0991***
	(0.0143)	(0.0451)	(0.0405)	(0.0179)	(0.0572)	(0.0534)	(0.0139)	(0.0417)	(0.0382)
Traffic	0.658***	0.667***	0.659***	0.674***	0.683***	0.674***	0.655***	0.664***	0.653***
	(0.0129)	(0.0417)	(0.0401)	(0.0128)	(0.0414)	(0.0400)	(0.0126)	(0.0406)	(0.0396)
N	5575	5301	5575	5575	5301	5575	5528	5254	5528
R^2	0.738	0.735	0.738	0.721	0.721	0.720	0.745	0.743	0.744
Districts	269	269	269	269	269	269	269	269	269

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 B.17: Alternative independent variables for PM_{10}

	Alternative population density			Log of population			Log of Employed per area		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil
logaltdensity	0.0281***	0.0424***	0.0268*						
	(0.00321)	(0.0106)	(0.0146)						
logpopulation				0.0519***	0.0715***	0.0408***			
				(0.00317)	(0.0186)	(0.0158)			
logempldensity							0.0367***	0.0815***	0.0330
							(0.00528)	(0.0221)	(0.0300)
Distance to CBD	0.000489*	0.000734	0.000446	-0.000851***	-0.00111	-0.000769	0.000339	0.00107	0.000256
	(0.000287)	(0.000882)	(0.000860)	(0.000259)	(0.000682)	(0.000708)	(0.000293)	(0.000948)	(0.000942)
Suburban	0.103***	0.0911***	0.103***	0.106***	0.0967***	0.106***	0.109***	0.0944***	0.110***
	(0.00899)	(0.0240)	(0.0251)	(0.00872)	(0.0231)	(0.0239)	(0.00910)	(0.0240)	(0.0257)
Urban	0.171***	0.142***	0.172***	0.176***	0.150***	0.180***	0.176***	0.136***	0.178***
	(0.0102)	(0.0303)	(0.0310)	(0.00991)	(0.0292)	(0.0285)	(0.0102)	(0.0310)	(0.0318)
Industrial	0.129***	0.122***	0.130***	0.121***	0.109***	0.124***	0.134***	0.131***	0.134***
	(0.0125)	(0.0362)	(0.0352)	(0.0131)	(0.0364)	(0.0367)	(0.0130)	(0.0379)	(0.0381)
Traffic	0.259***	0.259***	0.259***	0.261***	0.266***	0.261***	0.260***	0.260***	0.261***
	(0.00675)	(0.0186)	(0.0180)	(0.00660)	(0.0176)	(0.0172)	(0.00686)	(0.0193)	(0.0183)
N	4648	4407	4648	4648	4407	4648	4601	4360	4601
R^2	0.462	0.453	0.462	0.486	0.477	0.484	0.454	0.436	0.454
Districts	247	247	247	247	247	247	247	247	247

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 B.18: Alternative independent variables for $PM_{2.5}$

	Alternative population density			Log of population			Log of Employed per area		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil
logaltdensity	0.0111	0.0298	0.0201						
	(0.00757)	(0.0234)	(0.0248)						
logpopulation				0.0475***	0.0657	0.0324			
				(0.00742)	(0.0470)	(0.0246)			
logempldensity							0.00307	0.0482	-0.00198
							(0.0112)	(0.0452)	(0.0441)
Distance to CBD	0.000949	0.00127	0.00126	0.000356	-0.0000299	0.000420	0.000287	0.000942	0.000173
	(0.000652)	(0.00153)	(0.00151)	(0.000583)	(0.00118)	(0.00123)	(0.000628)	(0.00159)	(0.00143)
Suburban	0.169***	0.160***	0.165***	0.162***	0.159***	0.166***	0.186***	0.174***	0.188***
	(0.0227)	(0.0511)	(0.0480)	(0.0218)	(0.0476)	(0.0451)	(0.0226)	(0.0497)	(0.0475)
Urban	0.214***	0.182***	0.201***	0.190***	0.172***	0.203***	0.222***	0.176***	0.227***
	(0.0243)	(0.0539)	(0.0563)	(0.0233)	(0.0535)	(0.0494)	(0.0246)	(0.0565)	(0.0608)
Industrial	0.0685***	0.0703*	0.0664*	0.0555***	0.0557	0.0605	0.0659***	0.0698	0.0662
	(0.0202)	(0.0418)	(0.0397)	(0.0203)	(0.0418)	(0.0387)	(0.0204)	(0.0433)	(0.0406)
Traffic	0.115***	0.116***	0.117***	0.124***	0.123***	0.120***	0.115***	0.114***	0.114***
	(0.0163)	(0.0393)	(0.0387)	(0.0156)	(0.0362)	(0.0371)	(0.0176)	(0.0427)	(0.0424)
N	795	758	795	795	758	795	774	737	774
R^2	0.244	0.220	0.243	0.276	0.258	0.272	0.246	0.214	0.246
Districts	109	109	109	109	109	109	109	109	109

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 B.19: Alternative independent variables for O_3

	Alternative population density			Log of population			Log of Employed per area		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil	OLS	IV density 1910	IV Soil
logaltdensity	-0.0751***	-0.0358***	-0.0735***						
	(0.00428)	(0.0127)	(0.0185)						
logpopulation				-0.0667***	-0.0675**	-0.0467*			
				(0.00525)	(0.0296)	(0.0239)			
logempldensity							-0.130***	-0.0672***	-0.157***
							(0.00642)	(0.0239)	(0.0318)
Distance to CBD	0.00292***	0.00417***	0.00297***	0.00557***	0.00556***	0.00544***	0.00264***	0.00393***	0.00212**
	(0.000296)	(0.000970)	(0.000947)	(0.000319)	(0.000992)	(0.000960)	(0.000290)	(0.000999)	(0.000949)
Suburban	-0.165***	-0.167***	-0.165***	-0.180***	-0.171***	-0.183***	-0.162***	-0.166***	-0.156***
	(0.0111)	(0.0351)	(0.0342)	(0.0113)	(0.0346)	(0.0351)	(0.0113)	(0.0351)	(0.0350)
Urban	-0.222***	-0.246***	-0.223***	-0.266***	-0.257***	-0.270***	-0.218***	-0.242***	-0.205***
	(0.0119)	(0.0386)	(0.0396)	(0.0118)	(0.0392)	(0.0382)	(0.0120)	(0.0383)	(0.0401)
Industrial	-0.0516***	-0.0344	-0.0513*	-0.0392***	-0.0253	-0.0387	-0.0743***	-0.0439	-0.0825**
	(0.0117)	(0.0351)	(0.0310)	(0.0151)	(0.0387)	(0.0455)	(0.0124)	(0.0357)	(0.0340)
Traffic	-0.226***	-0.236***	-0.226***	-0.234***	-0.246***	-0.232***	-0.222***	-0.234***	-0.221***
	(0.0165)	(0.0410)	(0.0394)	(0.0177)	(0.0440)	(0.0432)	(0.0163)	(0.0402)	(0.0385)
N	3776	3588	3776	3776	3588	3776	3751	3563	3751
R^2	0.427	0.407	0.427	0.389	0.389	0.386	0.430	0.415	0.426
Districts	251	251	251	251	251	251	251	251	251

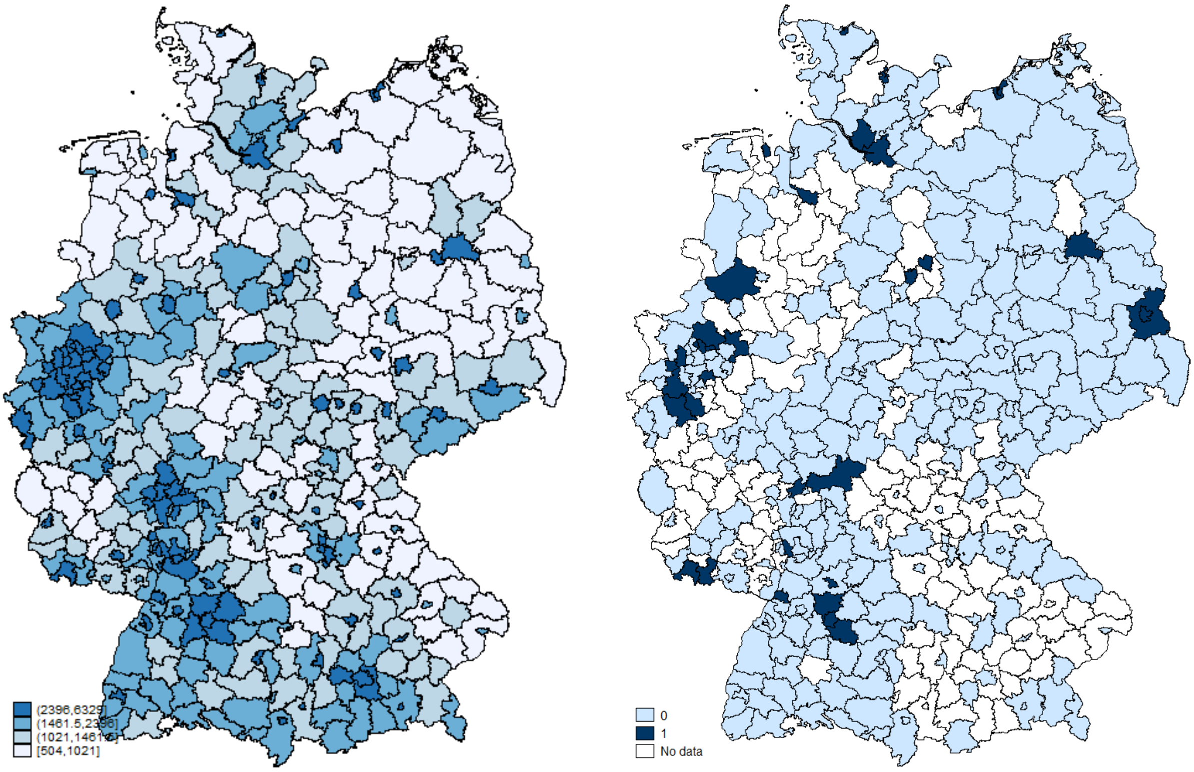
Standard errors in parantheses are clustered at labor market region - year (OLS) and labor market region (IV) level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.20: Regressions including an interaction term for population density*station type

	NO$_2$			PM$_{10}$			PM$_{2.5}$		
	(1) OLS	(2) IV Historic	(3) IV Soil	(4) OLS	(5) IV Historic	(6) IV Soil	(7) OLS	(8) IV Historic	(9) IV Soil
log(pop density)	0.255***	0.110*	0.235***	0.0548***	0.0899***	0.0145	0.0169	0.0579	0.00617
	(0.0159)	(0.0595)	(0.0725)	(0.00728)	(0.0292)	(0.0479)	(0.0170)	(0.0579)	(0.0595)
Industrial*Density	0.127***	0.239***	0.143*	0.0387**	0.00970	0.0692	0.0574*	0.0416	0.0631
	(0.0249)	(0.0823)	(0.0850)	(0.0195)	(0.0642)	(0.0650)	(0.0298)	(0.0600)	(0.0601)
Traffic*Density	0.0589**	0.192**	0.0774	0.0623***	0.0369	0.1000**	0.0412	0.0291	0.0501
	(0.0230)	(0.0768)	(0.0821)	(0.0113)	(0.0359)	(0.0454)	(0.0331)	(0.0821)	(0.0797)
Distance to CBD	-0.00380***	-0.00565***	-0.00402***	0.000607**	0.000786	0.000124	0.00111*	0.00131	0.000983
	(0.000469)	(0.00165)	(0.00149)	(0.000283)	(0.000868)	(0.000892)	(0.000657)	(0.00154)	(0.00146)
Suburban	0.357***	0.368***	0.361***	0.109***	0.0946***	0.115***	0.177***	0.167***	0.179***
	(0.0158)	(0.0482)	(0.0473)	(0.00909)	(0.0242)	(0.0276)	(0.0238)	(0.0538)	(0.0521)
Urban	0.524***	0.573***	0.531***	0.176***	0.147***	0.191***	0.218***	0.184***	0.225***
	(0.0196)	(0.0603)	(0.0630)	(0.0101)	(0.0304)	(0.0339)	(0.0257)	(0.0589)	(0.0646)
Industrial	-0.873***	-1.722***	-0.996	-0.157	0.0573	-0.388	-0.360	-0.237	-0.402
	(0.193)	(0.636)	(0.655)	(0.143)	(0.465)	(0.477)	(0.224)	(0.462)	(0.465)
Traffic	0.200	-0.841	0.0551	-0.230**	-0.0301	-0.525	-0.204	-0.111	-0.274
	(0.182)	(0.611)	(0.649)	(0.0908)	(0.286)	(0.358)	(0.262)	(0.657)	(0.638)
N	5575	5301	5575	4648	4407	4648	795	758	795
R^2	0.750	0.742	0.750	0.472	0.463	0.468	0.250	0.229	0.249
Districts	269	269	269	247	247	247	109	109	109

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



(a) Density in 2015

(b) Districts with coal-fired plants

Figure B.2: Population densities in 2015 and districts with coal-fired plants

Table B.21: IV regressions with historical population density including historical share of workers in industry

	NO_2		PM_{10}		$PM_{2.5}$		O_3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(pop density)	0.186***	0.170***	0.103***	0.0833***	0.0766	0.0356	-0.0969***	-0.0815***
	(0.0533)	(0.0550)	(0.0240)	(0.0291)	(0.0522)	(0.0488)	(0.0292)	(0.0316)
Distance to CBD	-0.00550***	-0.00578***	0.000972	0.000796	0.00146	0.00160	0.00371***	0.00415***
	(0.00153)	(0.00159)	(0.000850)	(0.000858)	(0.00161)	(0.00143)	(0.000960)	(0.00101)
Suburban	0.335***	0.333***	0.0880***	0.0858***	0.158***	0.127***	-0.168***	-0.165***
	(0.0459)	(0.0463)	(0.0231)	(0.0230)	(0.0506)	(0.0482)	(0.0337)	(0.0350)
Urban	0.537***	0.530***	0.141***	0.133***	0.173***	0.162***	-0.249***	-0.243***
	(0.0596)	(0.0596)	(0.0307)	(0.0302)	(0.0545)	(0.0487)	(0.0388)	(0.0400)
Industrial	0.0710*	0.0804*	0.130***	0.111***	0.0587	0.0634*	-0.0318	-0.0403
	(0.0390)	(0.0419)	(0.0356)	(0.0360)	(0.0427)	(0.0326)	(0.0322)	(0.0343)
Traffic	0.682***	0.674***	0.261***	0.261***	0.115***	0.140***	-0.267***	-0.237***
	(0.0428)	(0.0420)	(0.0188)	(0.0178)	(0.0399)	(0.0388)	(0.0383)	(0.0411)
Share employed in Ind.	1.190**		-0.0192		0.496		-1.588***	
	(0.543)		(0.280)		(0.580)		(0.388)	
Share workers in Ind.		0.141		0.234***		0.444***		-0.0448
		(0.129)		(0.0813)		(0.128)		(0.106)
N	5178	5178	4326	4326	749	749	3493	3493
R^2	0.749	0.744	0.459	0.476	0.228	0.299	0.446	0.412
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 B.22: Probability of transgression of annual threshold thresholds using a probit model

	NO_2			PM_{10}			$PM_{2.5}$		
	(1) Probit	(2) Probit Hist. IV	(3) Probit Soil IV	(4) Probit	(5) Probit Hist. IV	(6) Probit Soil IV	(7) Probit	(8) Probit Hist. IV	(9) Probit Soil IV
main									
log(pop density)	1.477***	1.228***	1.840***	0.464***	0.633***	0.333	-0.260	-0.677*	-0.435
	(0.218)	(0.252)	(0.279)	(0.105)	(0.128)	(0.264)	(0.249)	(0.392)	(0.562)
distance to CBD	0.0191*	0.0133	0.0222**	0.00627	0.00752	0.00424	0.00926	-0.00320	0.00695
	(0.0107)	(0.0108)	(0.0106)	(0.00451)	(0.00465)	(0.00594)	(0.0187)	(0.0186)	(0.0200)
suburban	0.129	0.0886	0.0289	0.811***	0.752***	0.820***	1.834***	1.617***	1.860***
	(0.486)	(0.463)	(0.532)	(0.150)	(0.145)	(0.154)	(0.475)	(0.484)	(0.480)
urban	0.792	0.763	0.501	1.052***	0.927***	1.094***	2.320***	2.202***	2.456***
	(0.497)	(0.475)	(0.575)	(0.183)	(0.184)	(0.197)	(0.658)	(0.692)	(0.768)
industrial	0.0418	0.127	-0.0390	1.184***	1.156***	1.177***			
	(0.260)	(0.281)	(0.273)	(0.191)	(0.188)	(0.193)			
traffic	2.505***	2.514***	2.415***	1.691***	1.727***	1.693***	0.378	0.406	0.345
	(0.250)	(0.247)	(0.261)	(0.145)	(0.144)	(0.144)	(0.468)	(0.426)	(0.452)
logpopdensity									
distance to CBD		-0.00448***	-0.0127***		-0.00551***	-0.0127***		-0.00174	-0.0148***
		(0.00142)	(0.00217)		(0.00134)	(0.00234)		(0.00267)	(0.00439)
suburban		0.0251	0.00850		0.0497	0.0409		0.0904	0.114
		(0.0535)	(0.0906)		(0.0526)	(0.0868)		(0.0948)	(0.171)
urban		0.0589	0.190**		0.0683	0.216***		0.260**	0.409**
		(0.0529)	(0.0787)		(0.0499)	(0.0739)		(0.102)	(0.167)
industrial		-0.0626	-0.122*		-0.0497	-0.143**			
		(0.0542)	(0.0695)		(0.0646)	(0.0700)			
traffic		0.0133	0.0477		-0.00894	0.0646**		0.0111	-0.0182
		(0.0273)	(0.0299)		(0.0273)	(0.0329)		(0.0557)	(0.0618)
N	5663	5383	5663	4817	4565	4817	701	650	701
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Districts	269	269	269	247	247	247	109	109	109
Marginal Effects	0.2006	0.2532	0.1330	0.0924	0.0331	0.1110	-0.0132	0.0003	-0.0030

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

Table B.23: Probability of transgressing thresholds by specific number of days using a probit model

	NO_2			PM_{10}			$PM_{2.5}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	>17	>14	>9	>34	>29	>24	>34	>29	>24
main									
log(pop density)	1.141***	1.219***	1.183***	0.377***	0.372***	0.266**	-0.0978	0.00227	-0.0188
	(0.435)	(0.442)	(0.408)	(0.129)	(0.114)	(0.107)	(0.142)	(0.138)	(0.142)
distance to CBD	0.00112	0.00692	0.00695	0.00647	0.00341	0.00180	0.00191	0.00356	0.00881
	(0.0321)	(0.0300)	(0.0230)	(0.00548)	(0.00495)	(0.00489)	(0.00819)	(0.00648)	(0.00942)
suburban	-0.254	-0.293	-0.250	0.303	0.362*	0.429**	1.264***	1.061***	1.059***
	(0.578)	(0.587)	(0.539)	(0.226)	(0.207)	(0.168)	(0.317)	(0.280)	(0.282)
urban				0.298	0.330	0.439**	1.318***	1.125***	1.224***
				(0.257)	(0.243)	(0.191)	(0.335)	(0.308)	(0.307)
industrial	-0.599	-0.636	-0.848**	0.953***	0.817***	0.804***	0.231	0.496**	0.303
	(0.416)	(0.421)	(0.407)	(0.250)	(0.252)	(0.250)	(0.191)	(0.213)	(0.188)
traffic				1.503***	1.341***	1.267***	0.642***	0.695***	0.542***
				(0.105)	(0.112)	(0.103)	(0.192)	(0.184)	(0.189)
N	2125	2125	2125	4817	4817	4817	795	791	791
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Districts	269	269	269	247	247	247	109	109	109
Marginal Effects	0.0528	0.0588	0.0788	0.0514	0.0643	0.0560	-0.0248	0.0005	-0.0032

Standard errors in parantheses are clustered at labor market region level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$