The economic cost of air pollution: Evidence from Europe*

Antoine Dechezleprêtre[†] Nicholas Rivers[‡] Balazs Stadler[§]

Version: November 2, 2020

Abstract

We provide the first evidence that air pollution causes economy-wide reductions in market economic activity. We combine satellite-based measures of air pollution with statistics on regional economic activity throughout the European Union since 2000. We use an instrumental variables approach, based on both thermal inversions and the direction of prevailing wind, to identify the causal impact of air pollution on economic activity. We estimate that a 1 μg/m^3 (10 percent) increase in fine particulate matter concentrations causes a 1.1% reduction in gross domestic product, with over 90 percent of this impact due to reductions in output per person and the remaining amount due to reductions in population. Our estimates suggest that the economic benefits of reducing air pollution are much larger than previously thought, and of similar magnitude to the benefits associated with reductions in mortality.

*We acknowledge helpful comments from Shardul Agrawala, Ioannis Tikoudis, Nils-Axel Braathen, Alexander Mackie, Matthew Neidell, Max Auffhammer, Maureen Cropper, Dave Donaldson, the US EPA, seminar participants at London School of Economics, École des Mines, University of British Columbia, McGill University, University of Alberta, University of Oldenburg, OECD, Chaire de Climat (Université de Paris Dauphine), and at the 2018 IZA Workshop on Environment and Labor Markets, 2018 Toulouse Workshop on Environmental Regulation and Industry Performance, 2018 Canadian Resource and Environmental Economics Workshop, and the 2018 World Congress on Environmental and Resource Economics. The authors gratefully acknowledge funding from the European Union (Contribution agreement N 07.0201/2018/792535/SUB/ENV.F1)

[†]London School of Economics and Organisation of Economic Co-operation and Development. A.Dechezlepretre@lse.ac.uk

[‡]Graduate School of Public and Internatinal Affairs and Institute of the Environment, University of Ottawa. nrivers@uottawa.ca

[§]Organisation of Economic Co-operation and Development. balazs.stadler@oecd.org

1 Introduction

Air pollution represents a major threat to human health in the 21st century. The World Health Organisation estimates that only 1 in 10 people globally live in areas where air pollution is at recommended levels, and that outdoor air pollution is responsible for at least 4.2 million deaths a year globally and over 400,000 in the EU alone. Air pollution dominates all other major causes of avoidable death including tobacco smoking, alcohol use, road accidents, and transmissible diseases such as AIDS, malaria, and tuberculosis. Since air pollution continues to rise at an alarming rate worldwide, especially in low- and middle-income countries, these numbers may grow even larger in the years to come (OECD, 2016; European Commission, 2013).

The consequences of air pollution on human health have led to the introduction of increasingly stringent environmental regulations around the world (Botta and Koźluk, 2014), but controversy remains over their appropriate stringency. Imposing environmental regulations is typically seen as a trade-off between generating non-market benefits to health or the natural environment but imposing costs on the economy, as resources may be redirected away from productive activities towards pollution control activities. Therefore, this debate is often framed in terms of "jobs versus the environment" (e.g. Morgenstern, Pizer, and Shih 2002). However, this framing ignores the potential that reductions in air pollution may lead to improved productivity, which itself can translate into greater economic output.

The objective of this paper is to inform this debate by estimating the causal impact of air pollution on economic activity, using data from across Europe. The results show that higher levels of air pollution, as measured by $PM_{2.5}$ concentration (small airborne particles with a diameter less than 2.5 microns, the pollutant with by far the largest estimated impacts on mortality and health outcomes), exert a substantial direct burden on the market economy, principally by reducing output per worker. This implies that reducing air pollution could yield large market economic dividends in addition to the well-established non-market benefits, and suggests that prior estimates of the benefits of pollution reduction are substantially too low.

In cost-benefit analyses of air pollution control policies, the benefits are typically vastly dominated by non-market impacts such as avoided deaths. In contrast, market benefits–such as reduced absenteeism at work–appear of second order importance in these evaluations. For example, the U.S. Environmental Protection Agency estimates that the benefits of the Clean Air Act Amendments over the period 1990-2020 amount to $12 trillion (in 2006 USD), with 85 percent of these benefits attributable to reductions in premature mortality (US EPA, 2011). Similarly, recent analysis by the OECD estimates that the total market costs of outdoor air pollution (including reduced agricultural yields, absenteeism at work and health expenditures) amount to 0.3 percent of global income in 2015 while the welfare costs from non-market impacts represent 6 percent of total income (OECD, 2016).

However, recent evidence is emerging that suggests poor air quality may cause direct reductions in economic activity because it negatively impacts cognitive or physical ability (Graff Zivin and Neidell (2018); Deryugina et al. (2019); Graff Zivin and Neidell (2012)). This literature–which we review in the following section–mostly focuses on observing the changes in individual productivity as well as on work absenteeism caused by concurrent exposure to poor air quality. This emerging literature points towards a consistent finding that air pollution negatively impacts the productivity of both low-skill and high-skill workers. However, it is difficult to draw a conclusion about the overall impact of poor air quality on the broader economy from these studies, which focus on idiosyncratic groups in particular locations, some of which in emerging economies with particularly high pollution levels (China, India). Another limitation is that prior studies make inferences about the effect of pollution on concurrent daily or even hourly economic performance, and for the most part have little to say about longer-run effects of pollution on productivity.

In this paper, we build on this literature by providing the first estimate of the causal impact of air pollution (measured by $PM_{2.5}$ concentration) on aggregate economic activity in a developed country context, using regional data from Europe for the period 2000-2015. We focus on the relationship between annual pollution and economic outcome measures, for the population at large, and thus get around both the concern about idiosyncratic populations as well as potential productivity displacement effects within a year. Our study is based on data from highly disaggregated European administrative regions (NUTS3 regions, similar to U.S. counties) between 2000 and 2015, and thus reflects the impact of pollution on developed countries in a contemporaneous period.

Estimating the causal effect of air pollution on economic outcomes at an aggregate level is challenging because of the potential for reverse causality. Not only might air pollution impact economic output and productivity (the effects we seek to measure), but economic activity clearly also affects pollution emissions through a number of potential channels. To circumvent this problem, we adopt an instrumental variables strategy, in which we use thermal inversions as well as wind direction as instruments, which generate quasi-random variation in pollution. Both of these instruments are strong, in that they predict pollution, are exogenous, in that they are not themselves caused by economic activity or pollution, and do not affect economic outcomes (conditional on weather) except through their effect on pollution.

The results show that air pollution adversely economic activity substantially. A 1 $\mu g/m^3$, or roughly 10 percent, increase in the average annual concentration of air pollution causes a 1.1% reduction in real gross domestic product in the average NUTS3 region. This implies that a 10 percent reduction in $PM_{2.5}$ concentration across Europe would increase European GDP by about €150 billion. On a per capita basis, this works out to about €300 per person. The impact of high pollution levels is heterogeneous across sectors, with the agriculture sector being the most severely affected. Our results also suggest that the marginal damage from pollution is falling in pollution concentration, which helps to reconcile findings from this study with those undertaken in high-pollution regions of the world, and also suggests that air pollution can affect economic outcomes even in relatively low-pollution regions.

We undertake a range of robustness tests, including using alternative instrumental variables, using pollution measures from a number of different sources, including spatial and temporal lags in our estimation, removing outliers, and weighting observations, and find that the sign, significance, and magnitude of our key results is preserved.

These findings can inform ex-ante cost-benefit evaluations of air pollution reduction policies. On the benefits side, a back-of-the-envelope calculation suggests that the market benefits of reducing air pollution uncovered in this study are of similar magnitude to the widely recognized non-market benefits from reduced mortality. This compares with relatively small abatement costs: a recent assessment by the European Commission of the cost of reducing $PM_{2.5}$ emissions by 25 percent in the European Union would be €1.2 billion annually (European Commission, 2013). Our estimates suggest that the economic benefits from such emissions reductions would be around two orders of magnitude greater. Therefore, much stronger air quality regulations could be warranted based on their previously underestimated economic benefits.

Simulations based on our statistical model show that the improvement in air quality between 2010 and 2020 required by the European Commission Ambient Air Quality Directives would increase European GDP by 1.25%, with some countries experiencing GDP growth of up to 3%. Environmental policies may also have contributed positively to economic growth in Europe in the recent period, and could further contribute to growth in the near future, as well as to the economic convergence between Western and Eastern European regions.

Our paper relates to the emerging literature which seeks to estimate the impact of air pollution on productivity and economic activity more generally. Alongside a number of studies that focus on the concurrent impact of pollution on individual outcomes (which we discuss in the following section), three papers use large-scale datasets that are representative of economy-wide economic activity. Probably closest to our paper is Fu et al. (2017), who use the near-universe of manufacturing plants in China and find that a $1\mu g/m^3$ increase in average annual $PM_{2.5}$ concentration reduces manufacturing sector productivity by 1.1%. Our paper complements Fu et al. (2017) by focusing on a developed country context where average pollution levels are much smaller than in China, and by estimating the impact on all sectors, rather than just manufacturing. The fact that our estimates are extremely similar to the ones reported by Fu et al. (2017) suggest that air pollution matters even at much lower average concentration levels. This finding is reinforced by our non-linear estimates, which show that an additional unit of $PM_{2.5}$ has a large impact at lower ambient pollution levels. Borgschulte et al. (2020) focus on pollution peaks in the U.S. caused by forest fires. They estimate that each day spent in a forest fire smoke plume causes a reduction in income of 0.04 percent over two years. Ignoring non-linearity, this result implies that continuous exposure to wildfire smoke would reduce income by a third. In contrast, our study focuses on the impact of exposure to pollution levels experienced on a daily basis by a typical resident of Europe. Isen et al. (2017) examine changes in lifetime income due to changes in exposure to pollution in infancy, and find that a 10 percent reduction in particulate matter during infancy causes a 1% increase in mean earnings later in life. Although the context is different, the magnitude is again similar to the findings reported in this paper.

Our paper is more generally related to the literature which seeks to understand the impact of environmental quality on economic activity. In particular, a growing literature empirically estimates the effect of temperature shocks on economic outcomes (see Auffhammer (2018) for a review). For example, Dell et al. (2009) document that in the year 2000, a 1°C warmer climate is associated with an 8.5 percent lower income per capita. Dell et al. (2012) find that economic growth is around 1 percentage point lower per additional °C. Burke et al. (2015) find evidence of a global non-linear relationship between temperature and economic production, and Deryugina and Hsiang (2014) find that the negative relationship between temperature and output is evident even in a rich country where adaptation opportunities (e.g., air conditioning) are presumably available.

The rest of the paper is organized as follows. Section 2 provides the background on the potential effects of pollution on economic outcomes. Section 3 describes our approach to estimating the causal effect of pollution on economic activity, including a discussion of our instrumental variable approach. Section 4 introduces the data. Section 5 provides the main results of our empirical analysis. Section 6 discusses the implications of our results, including by comparing our results to other studies, comparing the economic benefits of pollution reduction estimated in this study with estimates of mortality and morbidity benefits used in regulatory impact assessments, and by comparing our estimates of the benefits of pollution reduction to estimates of the cost of pollution reduction. Finally, Section 7 concludes.

2 Background

2.1 Conceptual framework

In this section, we provide a simple conceptual framework to illustrate the mechanisms through which pollution can impact economic output. The model is used to show how we measure the impacts of pollution on total economic output and to motivate the empirical analysis that follows. We also use the model to situate prior literature on the impact of air pollution on health and economic outcomes.

A representative firm in a closed economy has output given by $Y = Y(K, L, P)$, where Y is economic output, K is capital input, L is effective labor input, and P is pollution. We define y as per capita economic output, such that $Y = Ny$, where N is the population. Each of the N representative households has an endowment of time, and uses its income to finance consumption of the produced good. The total time endowment (t) of each household is specified as $t = h + s(P)$, where h is labor and where we use $s(P)$ to capture time periods in which the household is sick, and cannot work. Because the focus of this paper is not on optimal regulation of pollution, in this simple framework, we maintain pollution as an exogenous variable (see Graff Zivin and Neidell (2013) for a similar model in which pollution is treated as exogenous). The effective labour force available for work is $L = N(P)\phi(P)h$, where $\phi(P)$ reflects the impact of pollution on worker productivity, conditional on not being sick, and where we model the total population as a function of the level of pollution, to capture the idea that pollution can affect births, deaths, and migration. Given these assumptions, total economic output is given by:

$$Y = Y(K, N(P)\phi(P)\left[t - s(P)\right], P).$$

The impact of pollution on economic output is then given by:

$$\frac{d \log Y}{dP} = \psi \left[\frac{\partial \log N}{\partial P} - \theta \frac{\partial \log s}{\partial P} + \frac{\partial \log \phi}{\partial P} \right] + \frac{\partial \log Y}{\partial P}, \tag{1}$$

where ψ is the elasticity of output with respect to effective labor and $\theta = \frac{s}{t-s}$ is the benchmark ratio of sickness to labor supply.

In square brackets, the first term is the impact of pollution on total economic output as a result of changes in population. The second term is the impact of pollution on output as a result of changes in the number of hours worked, conditional on population. The third term is the effect of pollution on the productivity of the labour force. Finally, the last term on the right hand side (outside of the square brackets) captures the potential that air pollution directly affects economic output (not via its impact on labor supply or productivity). In the following subsection, we show that the literature suggests that pollution reduces economic output through all of the channels identified in this simple framework.[1]

In the empirical analysis that follows, we estimate the sign and magnitude of $\frac{d\log Y}{dP}$—the left hand side variable in (1). We decompose the change in economic output as a result of changes in pollution following our conceptual model. Unfortunately, we lack data on s as well as on any direct impact of pollution on output, and so our empirical decomposition does not parallel (1) exactly.

2.2 Prior literature

In this section, we provide evidence on the expected sign of each of the terms in (1): the impact of pollution on population, the impact of pollution on sickness and absenteeism, the impact of pollution on worker productivity, and the direct impact of pollution on output. It is not our intention to provide a systematic summary of the literature, but instead to highlight key results that help to provide a prior estimate of the key signs and magnitudes in (1).

2.2.1 Pollution and population ($\frac{d\log N}{dP}$)

It is widely recognized that air pollution imposes a substantial burden on human health (Graff Zivin and Neidell, 2013). Large cohort-based studies conducted by epidemiologists have provided evidence since at least 25 years ago that pollution by small airborne particles ($PM_{2.5}$) increases the rate of death (Dockery et al., 1993; Pope et al., 2002), especially through increases in respiratory and heart diseases. Calculations based on these and other studies suggest that ambient (outdoor) air pollution (especially $PM_{2.5}$) caused about 4.2 million deaths worldwide in 2015 (7.6% of all deaths), and was one of the leading causes of premature loss of life and loss of health (Cohen et al., 2017). Deryugina et al. (2019) estimates the short-run impact of air pollution on mortality, using a similar instrumental variables approach as in this paper, and finds that a $1\mu g/m^3$ increase in $PM_{2.5}$ results in a 0.2% contemporaneous increase in elderly mortality.

A substantial literature also finds evidence that pollution impacts birth outcomes. For example, Chay and Greenstone (2003) find that reductions in total suspended particulates (TSP, including both $PM_{2.5}$ as well as coarser particulates) caused reductions in infant mortality. They estimate that a 1-percent reduction in TSP reduced infant mortality by 0.35 percent in the early 1980s. Currie and Neidell (2005) find that reductions in PM_{10} and carbon monoxide (CO) in California both cause reductions in infant mortality. Jayachandran (2009) uses variation in exposure to smoke from the 1997 Indonesia forest fires to estimate "missing children" in downwind communities. She finds a large effect of exposure to forest fire smoke on infant mortality.

[1] For simplicity, we do not consider here the dynamics of capital accumulation, and thus ignore impacts of pollution on the capital stock.

Recent research also suggests that air pollution may impact migration. Chen et al. (2017) find great movement between provinces in China to avoid air pollution. Taken together, these studies suggest that air pollution likely reduces population in a region, by increasing deaths, reducing live births, and increasing net outmigration.

2.2.2 Pollution and absenteeism ($\frac{d \log s}{dP}$)

In addition to its effect on overall population, pollution has been found to affect sickness, and as a result, absenteeism. Ransom and Pope III (1992) provided early evidence on the relationship between outdoor pollution and absenteeism, by focusing on school attendance in Utah. They found that an increase in monthly PM_{10} of $100\mu g/m^3$ was associated with a 40% increase in absenteeism. Currie et al. (2009) report similar findings in Texas schools for CO.

Studies have also been conducted addressing absenteeism from work. For example, Holub et al. (2016) find that a $10\mu g/m^3$ increase in PM_{10} concentration results in a 1.6% increase in job absenteeism in Spain. Similarly, Hanna and Oliva (2015), Hansen and Selte (2000), and Aragon et al. (2017) show that increases in pollution reduce hours of work by a substantial magnitude. Interestingly, Aragon et al. (2017) finds that a key factor in explaining absenteeism from work, especially at moderate pollution levels, is the presence of dependents in the household (since, if a child is sick, a parent may have to stay home). Thus there may be a link between the school and work absenteeism outcomes.

2.2.3 Pollution and productivity ($\frac{d \log \phi}{dP}$)

In addition to causing substantial ill-health and mortality, air pollution is also believed to impair cognitive and physical function. Again $PM_{2.5}$ is of particular concern. When this pollutant is inhaled, the particles can enter deep into the lung and damage lung function. Additionally, they pass through the lung into the bloodstream, where they can affect the heart and brain function (Calderón-Garcidueñas et al., 2014; Du et al., 2016; Ranft et al., 2009). Because pollution affects physical and cognitive function, there is a clear pathway through which it could impact workplace productivity. Starting with Graff Zivin and Neidell (2012), a number of studies have investigated the link between productivity and other economic outcomes and elevated pollution. These studies have typically focused on groups of individuals for which productivity, or some similar measure, is directly observable and for whom tasks cannot easily be delayed or shifted in location.

Chang et al. (2016) examine the daily productivity of pear-packers at an indoor facility. They find that the number of boxes packed is reduced on days when air quality is poor. Adhvaryu et al. (2019) use data on hourly worker output at a garment manufacturing facility in India to show that increases in $PM_{2.5}$ concentrations cause reductions in worker productivity (measured by the number of garments sewn per hour). He et al. (2019) obtain data on worker-level output from two textile manufacturing facilities in China. They find that a sustained increase in $PM_{2.5}$ causes a reduction in worker output. Chang et al. (2019) show that the effect isn't limited to physical workers. They obtain a worker-level dataset from a Chinese call centre, and find that the number of calls handled by workers falls with increases in the air quality index, due to longer breaks at work taken by workers on polluted days.

Estimating the potential effect of pollution on high-skill workers is more challenging, because tasks are typically less routinized and can often be shifted in time and space. Nevertheless, there is some evidence that pollution also affects productivity in high-skill tasks. For example, Ebenstein et al. (2016) estimate the causal effect of poor air quality on student performance in standardized high-school examinations, and find that a $10\mu g/m^3$ increase in $PM_{2.5}$ concentration causes a 0.023% decline in exam scores. Archsmith et al. (2018) finds that the number of incorrect calls made by major-league baseball umpires increases by 2.6% when $PM_{2.5}$ increases by $10\mu g/m^3$, and Heyes et al. (2016) finds that a $7\mu g/m^3$ increase in $PM_{2.5}$ in New York causes a same-day fall of 12% in NYSE returns.

While it is difficult to generalize from these highly-specific tasks to the broader population, and while the magnitude of the measured impacts on these populations due to air pollution are quite varied, the emerging evidence points towards an increasingly consistent finding that air pollution impacts on-the-job outcomes, conditional on being at work. It is important to note that most of these studies focus on contemporaneous air quality and productivity, and so the estimates do not include any longer-run effects of pollution on productivity. An exception is Fu et al. (2017), who examines annual productivity, and He et al. (2019), who estimate productivity based on cumulative exposure over 25 days.

2.2.4 Direct impacts of pollution ($\frac{\partial Y}{\partial P}$)

In addition to impacts of pollution that are mediated through the labor market, air pollution may also have a direct impact on output. This is most likely in the agricultural or forestry sectors, where air pollution has the potential to damage crops or trees and thus cause reductions in yield.

A number of papers find that agricultural output is impacted by ambient pollution. Van Dingenen et al. (2009) use empirical dose-response relationships to estimate that current levels of pollution (primarily ozone) reduce global yields by 7-12% for wheat, 6-16% for soybean, and 3-4% for rice and maize. Avnery et al. (2011) report very similar results. Chameides et al. (1999) estimates that most crop yields in China are depressed by 5-30% as a result of suspended particulate matter, as this pollutant causes reductions in direct sunlight reaching plants, which is well known to depress yields. Schulze (1989) shows that deposition of air pollutant in soils affects soil acidity, and thus tree root development, long term growth rates, and tree health. Proctor et al.

(2018) estimates that pollution from the Pinatubo volcanic eruption substantially reduced agricultural yields (by about 5 to 10%, depending on crop). Outside of the agricultural sectors, Li et al. (2017) find that $PM_{2.5}$ pollution in China causes large losses in solar photovoltaic output (by 20% on an annual average basis in Eastern China) as it reduces direct radiation reaching solar panels.

2.2.5 Summary

These recent results, based on study populations around the world, suggest that air pollution affects population health and size, absenteeism, on-the-job productivity, and in some cases has direct impacts on output. While the evidence remains thin, the results also suggest that the on-the-job productivity of both low-skill and high-skill workers are affected by air pollution. Our aim in this paper is to tie these results together by examining overall impacts on economic performance due to high levels of air pollution.

Notably, estimates in our paper correspond to the aggregate effect of pollution, which could differ from prior studies for at least two reasons. First, prior studies could have been based on settings that were not representative of the entire economy. For example, many studies on pollution focus on particular activities because of data availability, rather than representativeness. Second, as a result of pollution, we expect that factors may reallocate away from negatively affected activities. Our approach aims to capture the net effect of pollution on economic activity, after most such adjustments have taken place (although it does not capture avoidance through moving across regions or shifting activities across years).

3 Empirical strategy

3.1 Econometric model

Consider a basic equation characterising the relationship between economic output and pollution concentration in region i in year t, which is the empirical analogue to equation (1):

$$\log(Y_{it}) = \beta_0 + \beta_1 P_{it} + \beta_2 f(W_{it}) + \eta_i + \gamma_t + \varepsilon_{it}, \tag{2}$$

where Y_{it} is a variable measuring economic output (GDP, GDP per capita, or gross value-added by sector), P_{it} is the average pollution concentration in region i in year t, $f(W_{it})$ is a flexible function that captures how economic output may be affected by weather (temperature, precipitation, etc), η_i are region fixed effects which capture any time-invariant differences between regions, such as differences in geography, γ_t are year fixed effects which account for changes in economic activity and pollution that occur across regions in the sample, and ε_{it} is a random disturbance term.

To sweep out the region fixed effects η_i and ensure that any persistent differences between regions, such as due to differences in geography, do not contribute to identification of the effect, and to address non-stationarity in the left hand side variable, we estimate Equation (2) in first differences:

$$\Delta \log(Y_{it}) = \beta_1 \Delta P_{it} + \beta_2 f(\Delta W_{it}) + \Delta \gamma_t + \Delta \varepsilon_{it}. \tag{3}$$

The coefficient β_1 can then be interpreted as the contemporaneous impact of pollution on GDP from a one-unit increase in the pollution concentration.

Our objective is to capture the causal effect of pollution on overall economic activity. This is not straightforward, because reverse causality is likely a major feature in this relationship. On the one hand, high levels of air pollution might increase absenteeism, mortality, and morbidity, and reduce workplace productivity, all of which contribute to reductions in overall economic activity, as in Equation (1). This is the effect we seek to investigate. However, on the other hand, changes in economic activity affect air pollution, through changes in technology, scale, preferences, regulations, trade, or other determinants of air pollution. As a consequence, a simple regression of economic outcomes on pollution, even controlling for other variables, will confound these two effects, and yield uninformative estimates of the effect of pollution on economic activity. In order to overcome the challenge associated with reverse causality, we require one or more variables that shift pollution quasi-randomly, and whose only effect on economic activity occurs via their effect on pollution. We adopt thermal inversions and wind direction as two such variables and carry out a two-stage estimation, in which we predict pollution in the first stage, based on observed prevalence of thermal inversions or the direction of wind, and in the second stage, estimate the effect of our predicted pollution measure on economic output. We explain the relevance of our instruments in the next sub-section.

The first stage of the model can be written as:

$$\Delta P_{it} = \alpha_1 \Delta TI_{it} + \alpha_2 \Delta WD_{it} + \alpha_3 \Delta f(W_{it}) + \lambda_t + \pi_{it}, \qquad \textit{(first stage)} \tag{4}$$

where TI_{it} is one or more measures of the frequency of thermal inversions in region i in year t, WD_{it} is one or more measures of the frequency of wind from different directions in year t and region i, and π_{it} is a disturbance term.

We then estimate the effect of our predicted pollution measure on economic output:

$$\Delta \log(Y_{it}) = \beta_1 \widehat{\Delta P}_{it} + \beta_2 \Delta f(W_{it}) + \gamma_t + \nu_{it}, \qquad \textit{(second stage)} \tag{5}$$

where ν_{it} is a random disturbance.

Both regressions are estimated in first differences, which implies that identification of the impact of pollution on economic activity is based on *within-region* differences in pollution. The regression equations also include time fixed effects (λ_t and γ_t respectively), to account for changes in economic activity and pollution that are common across regions, such as the 2008-09 economic downturn. In most of our results, we weight observations by population, such that results are not dominated by low-population regions (although we show that results are not particularly sensitive to weighting).

Note that we can also directly estimate the reduced form equation:

$$\Delta \log(Y_{it}) = \zeta_1 \Delta TI_{it} + \zeta_2 \Delta WD_{it} + \zeta_3 \Delta f(W_{it}) + \kappa_t + \phi_{it}, \qquad \textit{(reduced form)} \qquad (6)$$

which recovers the impact of our instrumental variables directly on economic activity (κ_t are time fixed effects).

It is important to note that the instrumental variable approach to estimating the effect of air pollution on economic activity also addresses the two other main sources of endogeneity, namely measurement error in air pollution–a feature of all studies on this topic (Graff Zivin and Neidell, 2013)–and omitted variables.

3.2 Instrumental variables

Our two-stage approach to estimating the effect of air pollution on economic activity requires instrumental variables that (1) affect pollution (i.e., are relevant instruments); (2) are not caused by pollution or economic activity (i.e., are exogenous and thus as good as randomly assigned); and (3) only affect the dependent variable through their effect on pollution, the endogenous variable (i.e., satisfy the exclusion restriction). We explain our choice of instrumental variables in the subsections below, and focus on how each of them satisfy these conditions.

3.2.1 Thermal inversions

The relationship between air temperature and pressure/altitude under normal atmospheric conditions is illustrated in Figure 1. Under normal conditions, air temperature decreases with altitude above the surface through the troposphere. At an altitude of roughly 11km above sea level (corresponding to 226 hPa), temperature reaches -56.5°C, and remains constant throughout the stratosphere before increasing towards the top of the atmosphere.

Thermal inversions occur in the lower troposphere, and represent a deviation from the normal monotonic relationship between air temperature and altitude/pressure. They form when a mass of cooler air becomes trapped below a warm mass of air. For example, the large-scale movement of

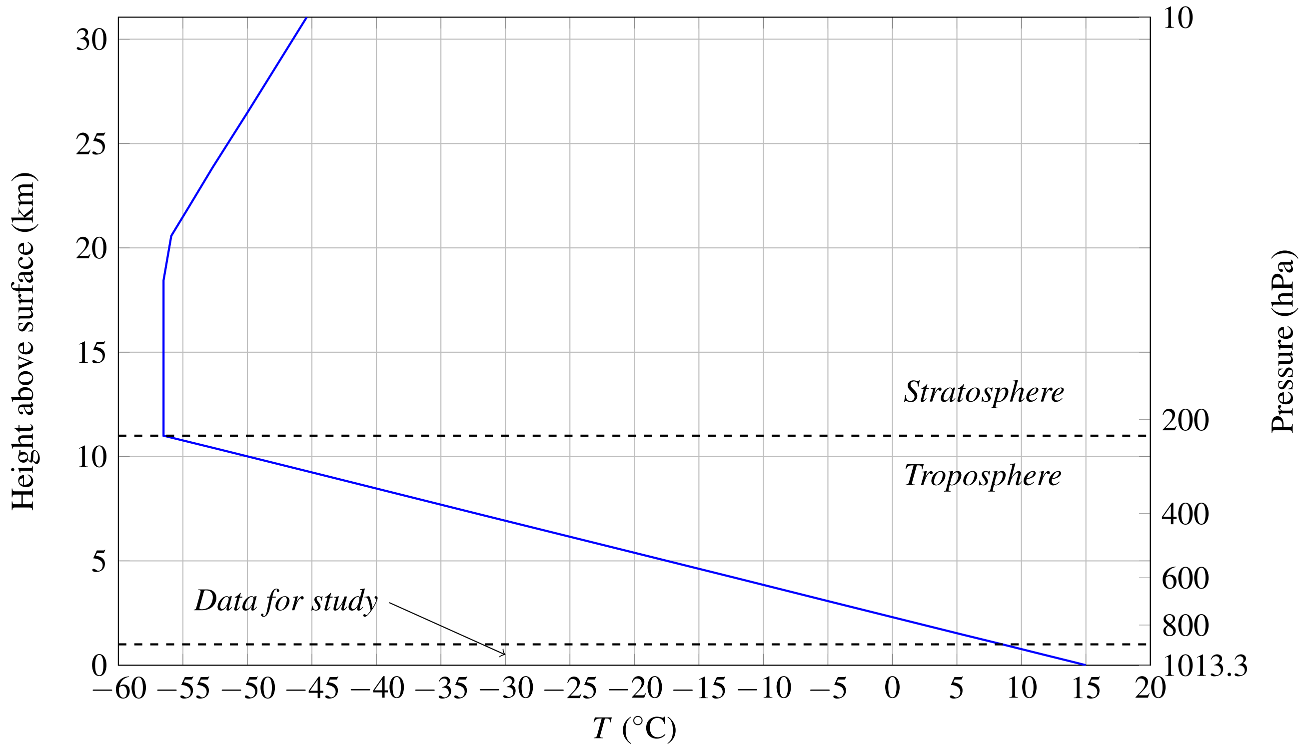

Figure 1: Typical atmospheric pressure and temperature. Under normal atmospheric conditions, temperature falls approximately linearly with height above surface until about 11km in elevation, where it reaches a temperature of about -56.5°C. Above that point, temperature is stable and then increasing until about 30km above surface at which point it decreases again until the top of the atmosphere. In this study, we retain temperatures within roughly 1,000m of surface—given by the lower dashed black line.

air masses throughout the atmosphere typically forms thermal inversions at its leading edge, as warm air masses pass over cooler air masses. Thermal inversions also form in winter at higher latitudes, as the low-angle sun heats the air higher in the atmosphere faster than the air at ground-level. Thermal inversions can also form as the surface cools overnight. Thermal inversions work with different mechanisms in winter and in summer. Summer inversions typically happen in the morning, whereas winter inversions typically take place in the afternoon, which also implies that they will have a different effect on pollution levels (Hicks et al., 2016).

Under normal atmospheric conditions, warm air at the surface is drawn upwards as a result of its lower density. This atmospheric ventilation can help to reduce pollution levels at the surface. During a thermal inversion, however, the inversion layer prevents the normal atmospheric ventilation from taking place, trapping polluted air at the surface. This effect is widely known, and has been documented in the scientific literature (Wallace and Kanaroglou, 2009; Gramsch et al., 2014). A similar strong relationship between thermal inversions and pollution is observed in other economics papers, such as Hicks et al. (2016), Chen et al. (2017), and Fu et al. (2017). In the results section, we formally demonstrate that air pollution, as measured by $PM_{2.5}$ concentration, increases significantly in the presence of thermal inversions.

Having established that thermal inversions are a *relevant* instrument, it remains to show that they are *as good as randomly assigned* (i.e., not caused by pollution or economic activity) and that they satisfy the *exclusion restriction* (i.e., only affect economic variables via their effect on pollution). Both of these must be shown by argument, rather than demonstrated using our data.

To demonstrate that thermal inversions are not caused by pollution or economic activity, we appeal to two branches of literature. First, in the climate economics literature, deviations in surface-level temperature from one year to the next within a region are typically assumed to be exogenous (e.g. Deryugina and Hsiang, 2017; Dell et al., 2012; Burke et al., 2015). Once we accept the exogeneity of the surface-level temperatures, the exogeneity of higher-altitude temperatures is easy to accept. Second, the atmospheric physics literature shows that aerosols can cause thermal inversions, by reflecting sunlight, but this happens only at extremely high levels of pollution, about 100 times larger than our sample average (50 times larger than our sample maximum) Rémy et al. (2015). Thus, in a European setting, reverse causality is not an issue.

To show that thermal inversions affect the economy only via pollution, it is important to remember that thermal inversions are an atmospheric phenomenon that takes place above ground level (where economic activity takes place). This fact should guarantee that thermal inversions satisfy the exclusion restriction. However, thermal inversions are linked with weather, which can potentially influence economic activity on the ground level (Dell et al., 2012; Burke et al., 2015). For example, thermal inversions often occur in winter, when surface temperatures are cooler. In order to rule out the potential correlation between inversions and economic conditions that occurs through weather, we carefully and flexibly control for on the ground weather conditions in all our regressions, as described below. These flexible controls for ground-level weather are given by the functions $f(W_{it})$ in Equations (4) and (5), and ensure that our instrument satisfies the exclusion restriction.

Several other papers have relied on similar arguments and used thermal inversions as an instrumental variable to understand the impact of pollution on behavioral and economic outcomes, such as Sager (2019), Arceo et al. (2016), Chen et al. (2017), and Bondy et al. (2020).

3.2.2 Wind direction

Airborne pollutants can be carried by the wind. This basic insight suggests that wind may be a *relevant* instrumental variable for predicting pollution concentrations. Indeed, there are a number of papers that use wind direction as an instrument for predicting pollution concentrations (Herrnstadt and Muehlegger, 2015; Ward, 2015; Deryugina et al., 2019; Bondy et al., 2020).

Unlike for thermal inversions, however, we do not begin with a clear idea of which wind direction may be associated with higher levels of pollution across all the geographical units in our study. To get around this, we use pre-sample data to establish which wind directions are associated with the highest levels of pollution in each region in our study sample (data is described in more detail in Section 4). Specifically, we gather daily data on wind direction and pollution concentration for the five-year period immediately preceding the period that we use for our main analysis. For each region in the data set, we calculate the average pollution concentrations associated with winds originating from each compass octant (for the pre-sample period). For example, Figure 2 shows the average $PM_{2.5}$ concentration associated with wind from each compass octant for a region in the Northwest of France (to the West of Paris). As shown in the Figure, for this region, in the pre-sample period, on days with winds blowing from the west, the pollution concentrations are on average lower than on days were the wind comes from the east. The most polluting wind directions for this region are the North-east and East, and the least polluting wind direction is the West.

We repeat this calculation for each of the 1,352 NUTS3 regions throughout Europe, again using pre-sample data to rank wind directions for each region according to the average pollution concentration. Figure 3 illustrates the unconditional relationship between daily wind direction and pollution concentrations across all of the regions covered by this study. We construct this figure using daily data for the five-year period immediately preceding the period that we use for our main analysis. For each region, we rank the compass directions from 1 to 8, where a ranking of 1 reflects the "cleanest" wind direction, and a ranking of 8 reflects the "dirtiest" wind direction, as described above.

The figure shows that the direction of the prevailing wind has an important impact on air pollution concentrations. On days when the wind originates from the "cleanest" direction, pollution concentrations are about half of what they are on days when the wind originates from the "dirtiest" direction. We use these observations based on pre-sample data to construct an instrument for pollution. In the main analysis, we use a variable to indicate the share of days in each year in which the wind originates from a "clean" direction (octants 1 through 3 in Figure 3) and another variable to indicate the share of days in which the wind originates from a "dirty" direction (octants 6 through 8).[2] As described above, the classification into clean and dirty directions is based on pre-sample data. In sensitivity analyses, we show that other reasonable choices for the wind instrument yield very similar results (for example, using only the dirtiest direction, the cleanest direction, or dummy variables for the number of days the wind originates from all eight directions). We formally show the relationship between our chosen instruments and pollution in the results section.

Wind patterns are caused by continental-scale movements in air masses, and as a result are

[2]Using the 4 cleanest and 4 dirtiest directions would produce variables that are colinear, so we drop the two middle directions.

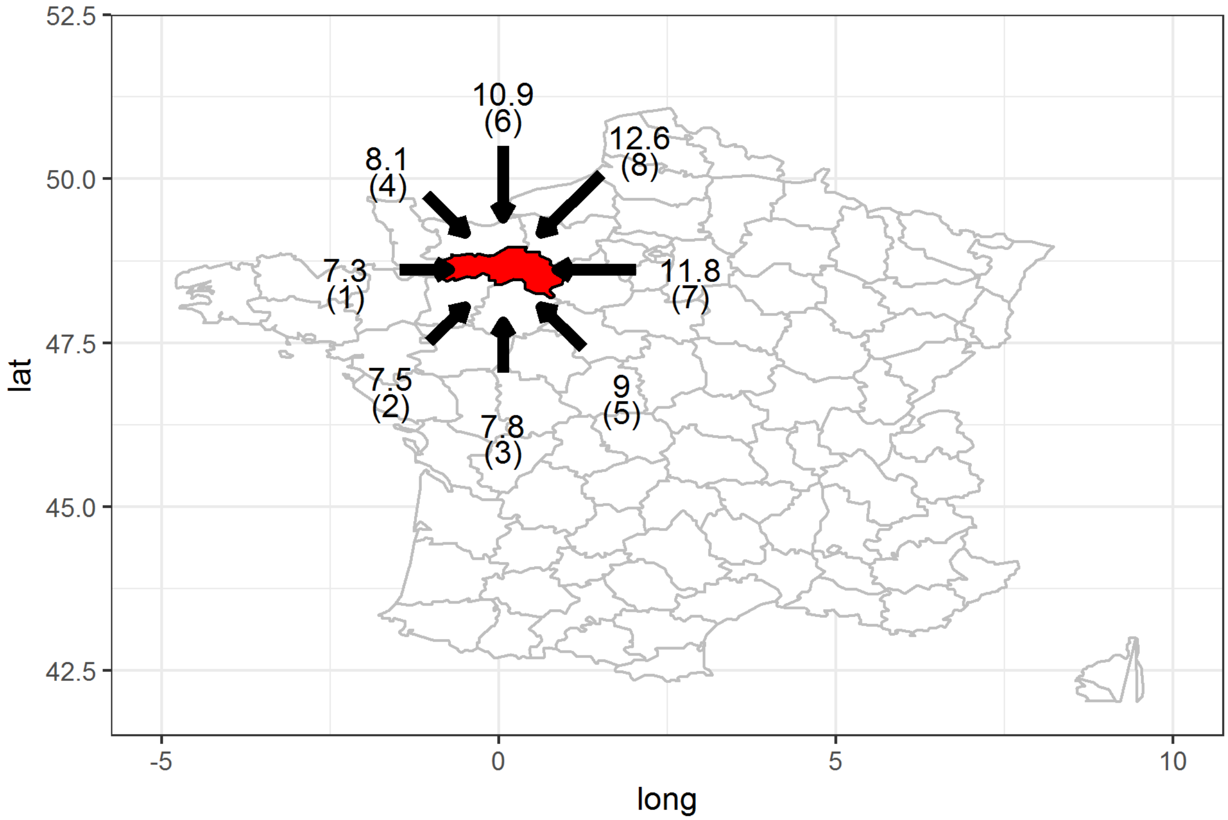

Figure 2: Illustration of method for categorizing wind directions using pre-sample data. The figure shows the average pre-sample (1995-1999) $PM_{2.5}$ concentration for each wind direction for a region in the Northwest of France, to the west of Paris (Départment Orne, NUTS FR253), highlighted in red. The length of the arrow corresponds to the average pollution concentration for wind blowing from the direction indicated by the arrow. Rankings of the wind directions from (1=) clean to (8=) dirty are given in brackets below the average $PM_{2.5}$ concentration.

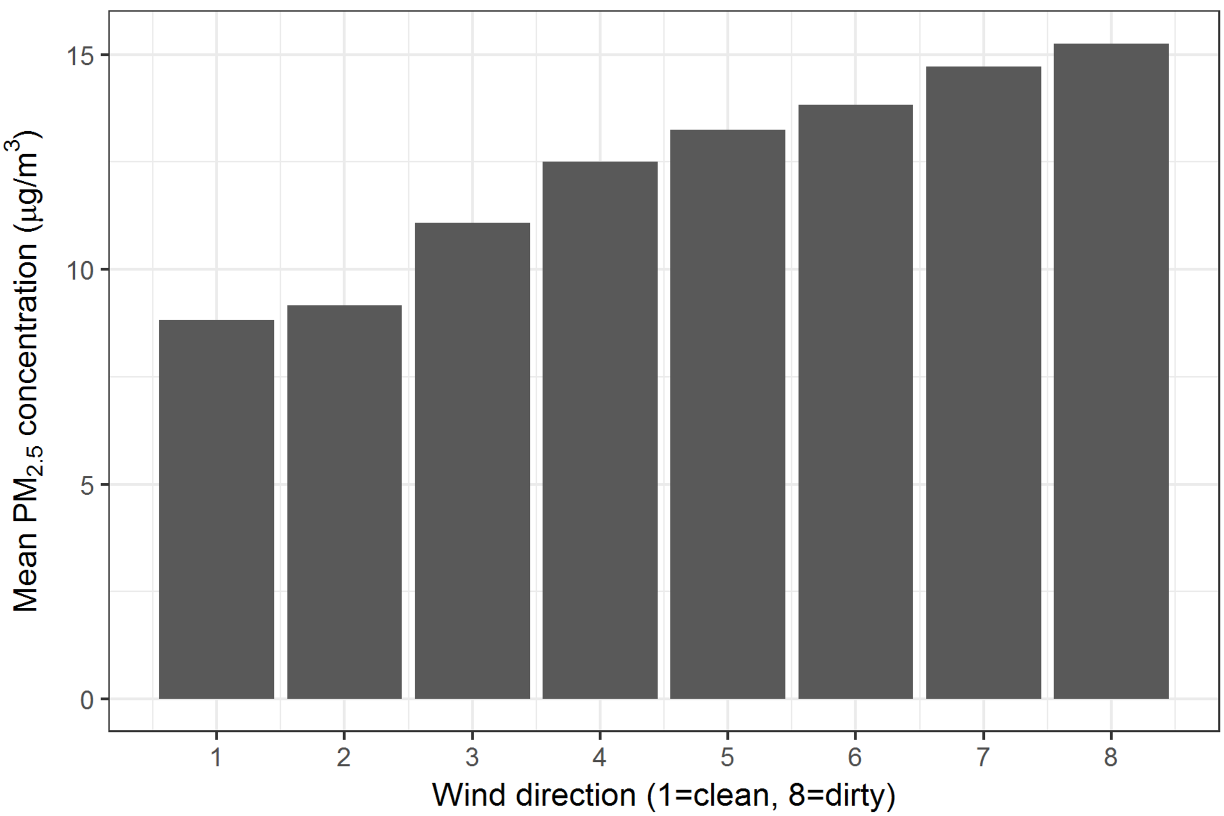

Figure 3: Relationship between daily average wind speed and wind direction and daily average $PM_{2.5}$ pollution concentrations. The ranking of wind direction is from 1(=cleanest) to 8(=dirtiest). Data used to construct this figure are drawn from the five-year period immediately preceding the main sample period, 1995-1999.

exogenous to the localized economic data that we employ. Wind direction is therefore an exogenous variable. Moreover, conditional on weather, the direction of wind satisfies the exclusion restriction. As described above, we condition very flexibly on both temperature and wind speed, as well as on other weather variables, such that the weather instrument is not confounded by correlation with weather. That is, there is no reason to believe that wind direction affects economic activity (conditioning on weather), save through its effect on pollution concentrations.

3.3 Weather covariates

As emphasized in the discussion above, we believe that our proposed instruments satisfy the exclusion restriction conditionally. That is, conditioning on weather covariates, the only pathway through which thermal inversions and wind direction affect economic activity is via their effect on pollution. To address possible confounding between weather and our instruments, we control very flexibly for a large variety of weather variables. Specifically, the function $f(W_{it})$ in equations (4) and (5) above includes a count of the number of days each year in which the average daily temperature falls into 20 temperature bins (that span the range of observed temperatures), a count of the number of days each year in which the daily average wind speed falls into one of 12 wind speed bins (defined using the Beaufort wind scale), a count of the number of days in each year in which precipitation falls into one of 20 exhaustive bins, second-degree polynomials in relative humidity and sea-level pressure, and interaction terms between all 20 temperature bins and both humidity and squared humidity. Additionally, we find very little sensitivity in our results to changes in the definitions of weather variables in regressions that we report later in the paper, suggesting that further changes in the set of weather controls are unlikely to influence the results.

4 Data

4.1 Air pollution data

The key endogenous independent variable in our model is air pollution. There are a a large number of potential air pollutants, and specific concern focuses on particulate matter, ground-level ozone, nitrogen oxides, and sulfur oxides.[3] Our analysis focuses on fine particulate matter, $PM_{2.5}$. There are two key reasons for this choice. First, $PM_{2.5}$ stands out as the pollutant with by far the largest estimated impacts on mortality and health outcomes. For this reason, the World Health Organi-

[3] Measures of these pollutants are used to construct the European Air quality Index, produced by the European Environment Agency: https://www.eea.europa.eu/themes/air/air-quality-index. The US EPA constructs its Air Quality Index using a similar range of pollutants (also including carbon monoxide): https://www3.epa.gov/airnow/aqi_brochure_02_14.pdf.

zation uses $PM_{2.5}$ concentration as an indicator of general population exposure to air pollution (World Health Organization, 2016). Similarly, most of the studies reviewed in the Section 2 use $PM_{2.5}$ as a proxy for air pollution more generally. Second, we are able to gather a comprehensive estimate of $PM_{2.5}$ concentrations covering the temporal and geographic scope required for our study.

Although we focus on this pollutant in our empirical application, it is important to emphasize that our empirical estimates of the impact of pollution on economic output likely confounds the effect of $PM_{2.5}$ with other air pollutants, since various air pollutants are typically correlated with one another, and since we mostly lack the data to control for other air pollutants (although we do use gridded SO_2 data in a multi-pollutant model). The various key ambient air pollutants share many sources in common—in particular they are all released as a by-product of combustion and industrial activity. As a result, our estimates should be taken as the effect of pollution on economic output, rather than the marginal effect of just $PM_{2.5}$ on economic output.[4]

It is useful to note that the IV approach to estimating the effect of air pollution on economic activity should help to address potential measurement error in air pollution, which is a feature of all studies on this topic (Graff Zivin and Neidell, 2013).

This section outlines the $PM_{2.5}$ data we use in our empirical analysis. In order to ensure that our results are not driven by a particular data set, we employ several alternative measures of air pollution in our analysis.

4.1.1 MERRA-2

Like a number of other papers, we make use of gridded air pollution data derived from a global reanalysis product. This has the advantage of providing complete geographic and temporal coverage for the period and units covered by our analysis. For our main specifications, we obtain air pollution data from NASA's MERRA-2 Aerosol product (MERRAero) (Buchard et al., 2017). MERRAero is a gridded aerosol and climate reanalysis product that produces a continuous estimate of aerosols since 1980 with complete global coverage. MERRAero produces an estimate of five different species of fine particulate matter, and we use the method of Buchard et al. (2016) to aggregate these into a consolidated estimate of $PM_{2.5}$ concentrations.[5] MERRA-2 estimates these particulate emissions by combining satellite measurements of aerosol optical depth (AOD),

[4]Most other research in this area is likewise unable to disentangle the individual effect of multiple pollutants. For examples, see Schlenker and Walker (2015) and Chang et al. (2018). In the robustness checks later in the paper we do introduce controls for co-pollutants and continue to find similar impacts of $PM_{2.5}$ on output.

[5]As in Buchard et al. (2016), we use the following calculation:

$$PM_{2.5} = [\mathrm{DUST}_{2.5}] + [\mathrm{SS}_{2.5}] + [BC] + 1.4 \times [OC] + 1.375 \times [\mathrm{SO}_4],$$

where SS is sea salt, BC is black carbon, and OC is organic carbon.

with estimates of particulate sources from an emissions inventory. The inputs are then assimilated using a global three-dimensional circulation model, including climate variables as well as aerosol transport and chemistry (this meterological model itself assimilates monitoring data on an extremely large number of relevant variables, such as surface temperature, wind, moisture, etc.). By combining satellite measures of aerosol optical depth (AOD) using an assimilation model based on well-understood physical and chemical dynamics, MERRAero achieves substantial improvements in fit compared to raw satellite AOD measures (Buchard et al., 2016).

The MERRA-2 pollution data contains hourly measures of aerosol concentration by species, which we aggregate as described above.[6] We retain the daily mean $PM_{2.5}$ concentration for each MERRA grid cell in the (longitude, latitude) range (-15,35) to (35,70).[7] We obtain air pollution data from January 1, 2000 to December 31, 2015. In our main empirical specification, we aggregate daily data up to an annual average. In addition, we obtain five years of daily observations of pre-sample data in order to rank wind directions according to average particulate concentrations, as described above. We choose MERRA-2 as our main source of air pollution data because it is the only data product with (1) complete spatial and temporal coverage for our data set, (2) high-resolution (daily) observations of air pollution, and (3) data available in the pre-sample period to construct our wind direction instrument.

4.1.2 Alternative measures of air pollution

We aim to ensure the results of our analysis are robust to alternative measures of air pollution. To do so, we conduct the analysis using alternative sources of air pollution data. The first alternative is based on Van Donkelaar et al. (2016). This product merges satellite air quality measurements with a geochemical transport model, and uses geographically-weighted regression based on surface air monitoring stations in order to obtain an improved match with surface air quality measures. Data is available at an annual basis on a very fine (0.01 or 0.1 degree) resolution grid. This data is widely-used. For example, the Lancet and World Health Organization use the data to produce the Global Burden of Disease report, and the OECD uses it to measure exposure to poor levels of air quality. Data is available at an annual frequency, and we obtain data from 2000 to 2015 for the entire region covered by our study.[8]

The second alternative data source is CAMS—the Copernicus Atmospheric Modeling Service.[9] This is an ensemble reanalysis, consisting of seven numerical air quality models developed

[6] We obtain our air pollution data from the M2T1NXAER files distributed by NASA. See: `https://disc.gsfc.nasa.gov/`.

[7] MERRA grid cells are 2/3 degree longitude and 1/2 degree latitute, or about 60 km by 60 km.

[8] Because Van Donkelaar et al. (2016) is low-frequency (annual), we continue to use MERRA-2 pre-sample data to construct our wind instrument.

[9] See `atmosphere.copernicus.eu`.

for Europe. These models incorporate meteorological data, satellite data, and ground-level air quality monitoring data into reanalysis products. Data is available from 2003-2015, and we use daily average data, which is merged to annual resolution.[10]

In addition, we obtain $PM_{2.5}$ pollution data from all ground-based monitoring stations in Europe from the European Environment Agency.[11] We impute average $PM_{2.5}$ concentrations by NUTS3 region as the annual average over all monitoring stations with the region. The ground monitoring network was very sparse at the beginning of our study period, and is gradually built out during the period covered by our analysis. The monitoring station data therefore contains a large number of missing observations. See Appendix A for additional details on the ground-based monitoring station data.

Although as a sensitivity test, we conduct our analysis with these three alternative air pollution measures, our preferred air pollution measure is the MERRA-2 product, which we use in our main results. We use MERRA-2 to obtain several of our other measures of atmospheric conditions, notably wind speed and direction and measures of thermal inversions. Because these are produced from a unified reanalysis, the MERRA-2 air pollution is physically consistent with other atmospheric measures from the same model. Moreover, it is the only product with full coverage over the analysis period that is available at a high temporal frequency (which we exploit above in categorizing wind directions in the pre-sample period).

4.2 Thermal inversions data

Thermal inversions data come from the MERRA-2 reanalysis.[12]. These files contain, among other variables, the air temperature at 42 vertical atmospheric levels (delineated by pressure) from 1000 hPa (near-surface) to 0.1 hPa (top of atmosphere). We obtain measures of daily mean air temperature at each grid cell, at each level from January 1, 2000 to December 31, 2015, and covering the bounding box defined by the coordinates (-15,35) to (35,70). We retain only observations within 1,000m above local surface level to focus on inversions that are most germane to ground-level pollution. There are 5 to 6 pressure levels in the MERRA-2 product in this region.

An inversion is a deviation from the normal monotonic declining relationship between air temperature and altitude. We operationalize this definition in three different ways, in order to ensure that our results are not sensitive to idiosyncratic choice of thermal inversion definition. A schematic overview of the manner in which we account for thermal inversions is given in Figure 4. First,

[10] Because CAMS data starts in 2003, we continue to use MERRA-2 pre-sample data to construct our wind instrument.

[11] See European Environment Agency Air Quality e-Reporting Database at `www.eea.europa/data-and-maps/data/aqereporting-2`

[12] We use the M2I3NPASM files distributed by NASA

we measure the presence of inversions at the lowest level of the atmosphere above the surface. Indexing atmospheric levels from $v = \{1, 2, \ldots 22\}$ with $v = 1$ representing the lowest atmospheric level above surface, we generate an indicator variable $\text{TI}^{LL} = \mathbb{1}(T_{v=2} - T_{v=1} > 0)$ that is equal to one if the temperature of the second layer is higher than that of the lowest layer (note that we drop indices for latitude, longitude, and time for clarity in this section; TI is short for Thermal inversion).[13] This measure of thermal inversions is closest to that adopted by Chen et al. (2017).

Second, we measure the presence of inversions at any atmospheric layer (below the 600 hPa layer and below 1,000m above surface level). We generate an indicator variable that is equal to one if there exists any pair of adjacent levels of the atmosphere, v and $v+1$, where the temperature of the upper layer is higher than that of the lower layer: $TI^A = \mathbb{1}(\exists v\,(T_{v+1} - T_v) > 0)$.

Third, we measure the presence of thermal inversions as the maximum deviation in temperature between all levels and the surface. To do so, we generate an indicator variable that is equal to one if there is any layer of the atmosphere (above 600hPa and below 1,000m above surface) in which the temperature is above the surface temperature: $TI^S = \mathbb{1}(\exists v\,(T_v - T_{v=1}) > 0)$.

In addition to measuring the presence of thermal inversions annually, it is also possible to measure the frequency of thermal inversions by season, since we observe temperature data at a high temporal frequency. Because the impact of thermal inversions may be seasonally heterogeneous, we separately measure the frequency of winter and summer thermal inversions using two instrumental variables for thermal inversions (in robustness checks, we show that alternative ways of accounting for thermal inversions generate very similar results).[14]

4.3 Weather and wind data

We obtain data on daily surface temperature, precipitation, and sea level pressure from the European Climate Assessment and Dataset.[15] This is a gridded product produced by amalgamation of all weather station data across Europe and interpolation (Haylock et al., 2008). The grid resolution is one quarter of a degree. We also observe daily surface temperature in the MERRA-2 data product, along with daily wind speed and direction as well as relative humidity.[16] All of our regressions produce identical conclusions whether using the European Climate Assessment or MERRA-2 surface temperature variables.

[13]It is important to note that in many cases, we do not observe temperature at the lowest pressure (1000 hPa) reported by MERRA because surface pressure is below 1000 hPa either because the land surface is elevated or due to a low pressure system. As a result, the set v is defined dynamically—in each grid cell in each day—with the index $v = 1$ always corresponding to the lowest pressure level above surface.

[14]We define summer as 16 April to 15 October and winter as 16 October to 15 April each year.

[15]See `http://www.ecad.eu/`.

[16]These variables are also derived from the M2I3NPASM files, and we use the same bounding box and temporal restrictions as for the thermal inversions data, above.

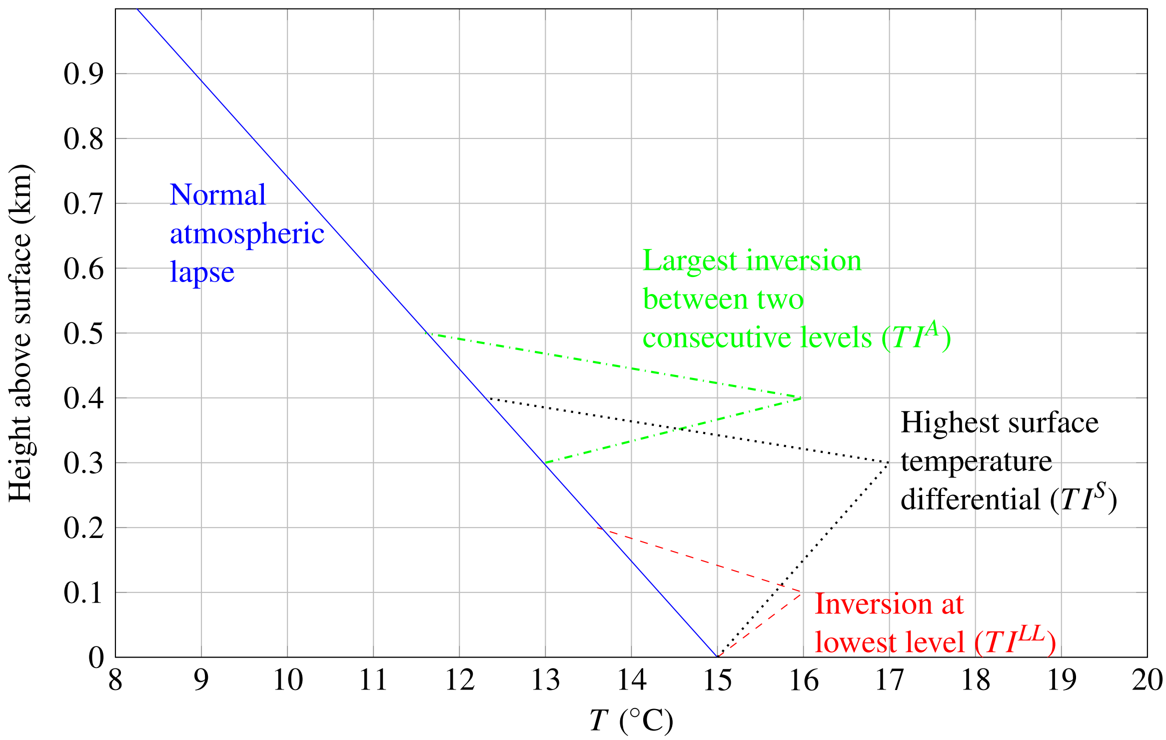

Figure 4: Schematic to define inversion variables used in the paper. The solid blue line shows the normal atmospheric lapse rate. Over the altitudes considered, temperature is monotonically declining with altitude under normal atmospheric conditions. The dashed red line shows our measure of inversions at the lowest level of the atmosphere. The dotted black line shows our second measure of inversions, which is the highest positive deviation between the surface and atmospheric temperatures. The dash-dotted green line shows our third measure of inversions, which is the largest positive deviation between any two adjacent atmospheric levels.

4.4 Economic outcomes data

We obtain data on economic outcomes from Eurostat's rural development database.[17] Our main indicator variable is gross domestic product at current prices by NUTS3 regions. We deflate this to real prices, and we refer to this measure of real economic output as Y_{it}, where i indexes NUTS3 regions and t indexes year. We also obtain data on gross value added by sector, deflate, and use this to measure economic outcomes at the sector level. We also obtain data from Eurostat on the annual population in each NUTS3 region. The source and construction of each of these variables are described in Appendix A.

4.5 Summary statistics

Figure 5 shows the average concentration of $PM_{2.5}$ over the period covered by our data. There is a substantial range in $PM_{2.5}$ concentrations, even on an annual level, with annual concentrations in some regions as low as $3.3\mu g/m^3$ in Northern Europe and over as $20\mu g/m^3$ in Eastern Europe and other regions. $PM_{2.5}$ concentrations are typically higher in Eastern Europe and the Mediterranean coast compared to the Atlantic coast and Scandinavia. It is important to note that the concentrations we report here are not exposure- or population-weighted, and thus significantly understate the typical pollution concentrations in urban areas. Figure 15 in Appendix B compares $PM_{2.5}$ concentrations from the gridded reanalysis product that we use in the analysis to measures from ground-based monitoring stations. While the two measures are highly correlated, the level of $PM_{2.5}$ concentration recorded by the ground-based monitoring stations—which are typically situated in urban regions with higher than average pollution—is about double the average level in the gridded reanalysis data.

Figure 6 shows the trends in the key instrumental variables as well as the endogenous variable, $PM_{2.5}$ concentrations, averaged over all of the regions. It is clear that there is no substantial upwards or downwards trend in the instrumental variables, but that there is some year-to-year variability in these variables, even at an aggregate level.[18] In addition, some correlation between thermal inversion strength and $PM_{2.5}$ concentrations is visually apparent from the figure—the years 2003 and 2011 stand out as the highest years for thermal inversions as well as the highest years for $PM_{2.5}$ concentrations in the figure.

Figures 8 and 9 in Appendix B show the identifying variation in two of the instrumental variables—thermal inversions and wind direction. Both figures show a histogram of these variables after first differencing and conditioning on year fixed effects. The figures show that even after

[17] The data is available at: `http://ec.europa.eu/eurostat/web/rural-development/data`.

[18] Note that identification of the econometric model is based on within-region variation in (instrumented) pollution and controls for time fixed-effects, so this figure does not show the variation that is used to identify the effect in the paper. Instead, it is provided as a summary of the data.

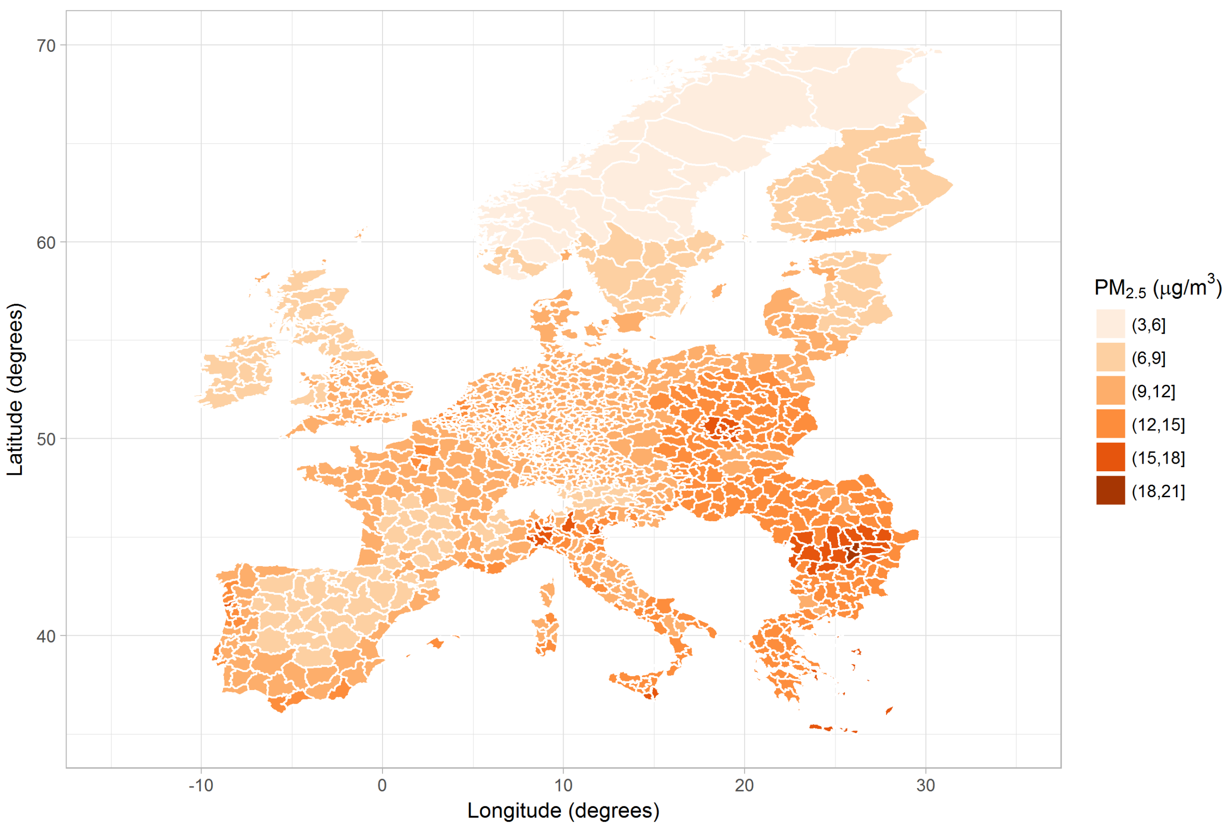

Figure 5: Map of average 2000-2015 $PM_{2.5}$ concentrations in economic regions used in the study.

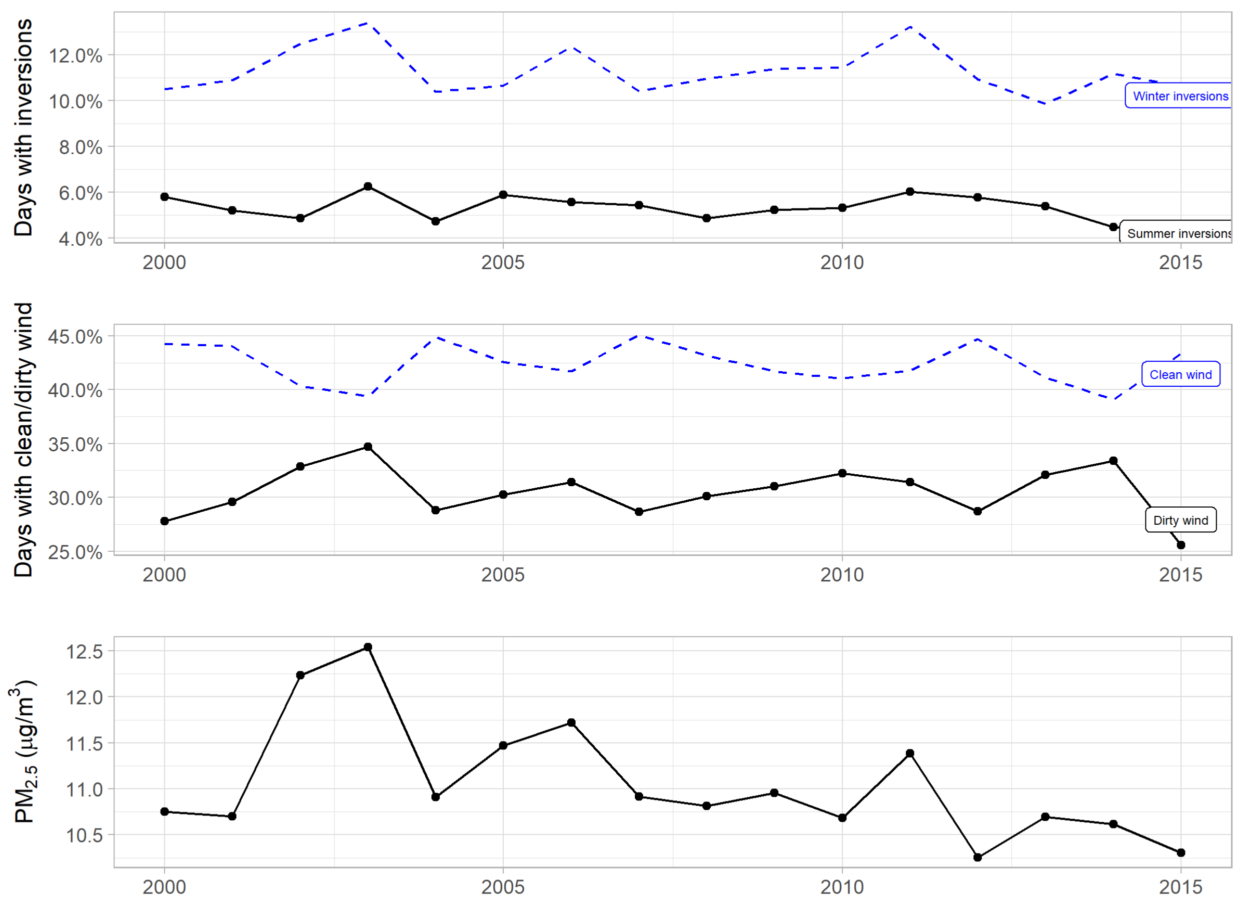

Figure 6: Annual trends in key independent variables.The top panel shows the variation in inversions over time, averaged across all of the regions in the data. The inversion measure captures low-level inversions (see text) in which the second atmospheric level is warmer than the lowest level. Inversions in both summer and winter periods are shown separately. The middle panel shows the proportion of days with winds from a "dirty" direction and "clean" direction (see text), averaged over all regions in the data. The bottom panel shows particulate matter concentrations averaged over all the regions in the data.

removing these fixed effects, there remains a substantial amount of variation in these variables. For example, it is normal to observe 5% or ($365\times0.05=$)18 days more or less of thermal inversions than the average (i.e., the standard deviation in the share of days with inversions is approximately 5%). A somewhat larger amount of variation exists for the wind instruments. It is evident there there exists considerable variation in these instruments even after controlling for geographic and time fixed effects, which we leverage for identification in our regressions. Appendix A also shows the distribution of geographic variance in the instrumental variables. The figure shows that wind direction is most variable throughout central Europe, such that these regions will contribute most to the identification of the effect of the impact of pollution on GDP using this instrument. In contrast, Northern Europe experiences the most variation in inversions from year-to-year, such that these regions will contribute most to identification of the impact of pollution on GDP using the inversion instrumental.

Summary statistics for all key variables in the data are provided in Appendix A.

5 Results

5.1 Main results

5.1.1 Ordinary least squares results

For completeness, we begin with results from an ordinary least squares specification without instrumental variables, i.e., equation (3). Results are given in Table 1. Recall that all regression coefficients are weighted by the population in each NUTS3 region to be representative of the average inhabitant in Europe rather than the average region. Column 1 shows the results estimated using a log-linear specification (log of GDP regressed on units of $PM_{2.5}$ pollution), while column 2 shows the results estimated in log-log form (log of GDP regressed on log of $PM_{2.5}$ pollution, which allows the results to be interpreted as an elasticity). Regressions are estimated in first differences and include year fixed effects. In all regressions, we condition flexibly on ground-level weather: we include counts of the number of days each year in each of 20 exhaustive temperature bins, counts of the number of days each year in each of 12 wind speed bins (corresponding to levels of the Beaufort scale), counts of the number of days in each year in each of 20 exhaustive precipitation bins, and second-degree polynomials in relative humidity and air pressure as well as interactions between temperature and humidity. Both regression coefficients indicate a statistically significant positive relationship between particulate matter and economic output. However, as explained above, these regression coefficients do not estimate the causal effect of pollution on economic output in which we are interested. Instead, they confound the impact of economic output on pollution with the impact of pollution on economic output, failing to identify either. To identify

Table 1: OLS estimation: association of PM2.5 and GDP.

	(1) ln(GDP)	(2) ln(GDP)
$PM_{2.5}$	0.0026 ***	
	(0.0004)	
$\ln(PM_{2.5})$		0.0192 ***
		(0.0049)
Observations	17099	17099

*Notes:** p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first difference and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

the causal effect of pollution on economic output, we instead turn to our instrumental variables approach.

5.1.2 First-stage results: The effect of wind direction and thermal inversions on pollution

Table 2 reports the results of estimating Equation (4), which is the first stage in our two-stage approach to estimating the effect of pollution on economic output. In this stage, we estimate the impact of the instrumental variables on pollution concentrations, after conditioning on weather covariates. In each case, we flexibly condition on weather as described above. Each regression also includes year fixed effects, while unit fixed effects are swept out due to the first-differencing approach, such that identification is from within-region variability. In the baseline regressions, we define the inversion instrument as the share of days in a year in which thermal inversions are observed, where inversions are defined as a positive vertical temperature gradient between the first two atmospheric layers (the variable TI^{LL}, defined earlier). We separately count winter and summer inversions, as described above, so there are two instrumental variables for thermal inversions. We define the wind direction instrument as the share of days in a year in which the wind in a region originates from one of the three "dirtiest" ranked compass octants and the share of days in a year in which the wind in a region originates from one of the three "cleanest" ranked compass

Table 2: First stage results: instruments effect on PM2.5.

	(1)	(2)	(3)
	$PM_{2.5}$	$PM_{2.5}$	$PM_{2.5}$
Summer inversions (low)	7.024 ***		5.913 ***
	(1.276)		(1.291)
Winter inversions (low)	-0.075		0.394
	(0.627)		(0.662)
Clean winds		-2.460 ***	-2.334 ***
		(0.264)	(0.283)
Polluting winds		1.562 ***	1.396 ***
		(0.285)	(0.269)
Observations	17099	17099	17099
Adjusted R^2	0.213	0.231	0.244

*Notes:** p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first difference and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. The instruments are defined as follows: inversions are the share of days in a year when an inversion is observed at the lowest atmospheric level, with summer and winter inversions separately counted. Clean (dirty) winds are the share of days in a year where the wind originates from one of the three cleanest (dirtiest) octants, based on pre-sample data. Regression coefficients weighted by each region's population.

octants (with ranking based on pre-sample data) as described earlier, so there are two instrumental variables for wind direction. For the wind instrument, the coefficients on clean and dirty winds can be interpreted as the impact on pollution relative to the two middle-ranked compass octants (neither clean nor dirty), while for the inversion instrument, the coefficients can be interpreted as the impact on pollution relative to a non-inversion day.[19] We focus here on log-linear results, but the results using a log-log specification deliver coefficients with nearly identical magnitudes and significance (see Appendix B.3.9 Table 23).

We use two sets of instruments—wind direction and thermal inversions—to generate exogenous variation in pollution. The instruments are not highly correlated with one another. For example, the correlation between the share of days where wind originates from one of the three dirtiest wind directions and the share of days with either summer or winter inversions is about 0.15 (after first differencing and controlling for year fixed effects), which suggests that they provide two sepa-

[19]We test the robustness of the results to alternative instrumental variables later in the paper. For detailed robustness checks see Appendix B.3 Tables 14 and 15.

rate sources of identification. Clean winds are similarly relatively uncorrelated with the inversion instruments.[20]

Column (1) of Table 2 shows the result when only thermal inversions are used as an instrument. Column (2) uses instead wind direction as an instrument, and column (3) uses both instruments simultaneously. The results suggest that both of our chosen instruments have a strong impact in the predicted direction on pollution concentrations. Specifically, summer thermal inversions cause a 6 to 7μg/m^3 increase in pollution concentrations. In contrast, we do not find a significant effect of winter inversions on pollution concentrations, consistent with winter inversions happening early in the morning and dissipating quickly as the sun starts shining. Likewise, changes in the direction of the wind cause substantial impacts on pollution concentrations. Winds from the "clean" directions cause a 2.4μg/m^3 reduction in pollution concentrations, and winds from "dirty" directions cause a 1.5μg/m^3 increase in pollution concentrations, relative to winds from "neutral" (middle-ranked) directions. Both of these effects are highly statistically significant and very large, suggesting that both chosen instruments are relevant. An F-test on the excluded instruments produces a value of 24 to 160, depending on which instruments are included, much higher than the weak instrument threshold of 10 normally adopted as a rule of thumb (Angrist and Pischke, 2008), again confirming the relevance of our selected instruments (see Table 3).[21]

5.1.3 Second-stage results: The effect of pollution on economic output

Table 3 presents the main results of the paper, which correspond to estimating equation (5) in which we regress economic activity on instrumented pollution and controls, with coefficients weighted by each region's population.[22] The organization of the table follows from Table 2 above. Specifically, column (1) uses thermal inversions as the instrument set, column (2) uses wind direction, and column (3) uses both sets of instruments simultaneously. The coefficient on instrumented pollution shows that a 1μg/m^3 increase in pollution concentrations causes a 1.1% to 1.2% reduction in economic activity depending on which instrument is used. In our preferred specification where both instruments are used (column 3), the marginal effect of a 1μg/m^3 increase in pollution concentrations on GDP is -1.16%. The effect is strongly statistically significant irrespective of which instrument, or combination of instruments, is used. The fact that we obtain similar results from two unrelated instruments lends substantial credibility to our results. The impact is also considerable in magnitude: an increase in $PM_{2.5}$ concentrations of 1μg/m^3 roughly corresponds to a 10% increase, suggesting the elasticity of GDP to pollution concentration is about -0.1. We present

[20] Appendix A.4 Table 9 provides a complete pairwise correlation matrix for the instrumental variables.

[21] The weak ID test reported throughout the paper is the Kleibergen-Paap rk Wald F statistic.

[22] Reduced form results showing the impact of wind direction and inversions on GDP are available in Table 10 in the Appendix.

log-log results in Appendix B.3.9 Table 24, and show that these deliver nearly identical results in magnitude and significance.

It is notable from Table 3 that both the wind direction and thermal inversion instruments suggest an economically and statistically significant negative effect of pollution on economic output. The magnitude of these effects differs by around 20%, and it is worth pointing out that there are two potential reasons for differences in magnitude of these effects. First, the instrumental variables estimator recovers the local treatment effect for "compliers": entities for which the instrument changes predicted pollution. In our case, the thermal inversion instrument changes pollution only in regions/years where there is variation in the share of days with thermal inversions from one year to the next (and similarly for the wind direction instrument). Appendix Figures 10 to 13 show that regions with variation in (summer and winter) inversions are different than those with variation in wind direction. In particular, Northern Europe experiences more substantial year-on-year variation in inversions, while Central Europe experiences higher year-on-year variation in wind direction. Thus the wind direction instrument will be more heavily weighted by Central European regions, and the thermal inversion instrument will be more heavily weighted by Northern European regions. Second, as discussed above, we proxy air quality by $PM_{2.5}$ concentrations, although other air pollutants also affect air quality. It is possible that wind and thermal inversions carry different mixes of co-pollutants, and thus the reported effect of $PM_{2.5}$ is associated with different mixes of pollutants using the two instruments. Despite these factors, the two instruments do recover a very similar point estimate for the effect of pollution on output.

In Equation (1) we show how we can decompose the effect of pollution on economic activity into four terms, reflecting the impact of pollution on population, presence at work, productivity, as well as a direct impact on output. Here, we use that framework to decompose the main results to better capture the mechanisms underpinning the effects we find. Unfortunately, we do not observe presence at work (s in Equation (1)) and cannot separately distinguish the direct impact of pollution on output ($\frac{\partial Y}{\partial P}$ in equation (1)) and so we are only able to decompose our results into two components—the effect of pollution on population and the effect of pollution on output per capita. We refer to this latter term as the impact of pollution on productivity, but it important to note that it is the joint effect of changes in work attendance, changes in work productivity conditional on attendance, as well as any direct impact of pollution on output, such as in the agricultural sector.[23] Table 4 reports the results. The result for the total effect of pollution on economic output is -

[23]Specifically, we set up an identity to decompose the results: $\mathrm{Y} \equiv \mathrm{Population} \times \mathrm{Y}/\mathrm{Population}$. We take logs and write the total derivative with respect to pollution as:

$$\frac{d\ln(\mathrm{Y})}{dP} = \frac{d\ln(\mathrm{Population})}{dP} + \frac{d\ln(\mathrm{Y}/\mathrm{Population})}{dP}.$$

We then estimate each of the terms on the right hand side in a separate regression, using both the wind and inversion instruments.

Table 3: IV estimations of the economic effect of PM2.5.

	(1) ln(GDP)	(2) ln(GDP)	(3) ln(GDP)
$PM_{2.5}$	-0.0106 **	-0.0121 ***	-0.0116 ***
	(0.0046)	(0.0023)	(0.0021)
Observations	17099	17099	17099
Weak id. stat.	23.98	160.2	107.0
Hansen J stat. p-value	0.751	0.0480	0.242

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first difference, and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. The first column uses the share of days in the year with (summer and winter) thermal inversions as instrument, the second column uses the share of days in the year with clean and dirty winds as an instrument, and the third column uses both sets of instruments. Regression coefficients weighted by each region's population.

0.0116 as in Table 3. Column (2) shows the effect of pollution on (labour) productivity (output per capita). The estimate is -0.0111, so an increase in pollution by $1\mu g/m^3$ decreases productivity by 1.1%. These point estimates imply that about ($\frac{-0.0111}{-0.0116}$ =)96% of the total effect of pollution on economic output is due to reduced productivity, with the remainder due to reduced population. This is confirmed in column 4 of the table, which shows that an increase in $PM_{2.5}$ concentration by $1\mu g/m^3$ causes a reduction in population by 0.05%. Column (3) of the table instead determines the effect of productivity by normalizing output by the working age population, and recovers a similar (but slightly smaller) impact on productivity. In each case, the main impacts of pollution on economic output occur as a result of reduced productivity, not reduced population.

5.2 Robustness checks

We conduct a large number of robustness checks to ensure that similar results are delivered by alternative choice of model and data. We report detailed robustness check results in Appendix B, and highlight key results and motivation for robustness checks here. Table 5 summarizes the key results from our robustness checks, in each case reporting the coefficient on instrumented pollution on output in our two-stage regression.

Weights. We re-estimate the model, weighting the observations by GDP rather than population (as in the baseline) and with unweighted observations. Weighting does not substantially affect our

Table 4: Decomposition of PM2.5's effect into effect on productivity and population.

	(1) ln(GDP)	(2) ln(GDP/pop.)	(3) ln(GDP/work pop.)	(4) ln(Population)
$PM_{2.5}$	-0.0116 ***	-0.0111 ***	-0.0103 ***	-0.0005 *
	(0.0021)	(0.0021)	(0.0021)	(0.0003)
Observations	17099	17099	17099	16204
Weak id. stat.	107.0	107.0	107.0	100.3
Hansen J stat. p-value	0.242	0.0756	0.0754	0.0000183

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Each column reports instrumental variable estimates of variables in each column on pollution. Instrumental variables are the share of days in a year with winter and summer thermal inversions, and the share of days in a year with clean and dirty winds. Regression coefficients weighted by each region's population.

results. Neither estimate is statistically different from the baseline, but they do suggest that the impact of air pollution on economic activity is slightly larger in small regions (as measured by GDP or population) than in larger ones. In our policy simulation results later in the paper, we use the GDP-weighted estimates as they provide the cleanest impact on aggregate European GDP, but note that using population-weighted estimates provide very similar results.

Dropping outliers. To ensure that our results are not driven by extreme values, we re-run the regression, but this time dropping extreme values. Specifically, we remove the observations below the 1^{st} (or 5^{th}) and above the 99^{th} (or 95^{th}) percentile of pollution levels. In both cases, we continue to obtain similar regression results, with our coefficient estimate increasing somewhat in absolute magnitude when we drop a larger number of observations.

Spatial lags and autocorrelation. Results are attenuated somewhat, but remain large and precisely estimated, when adding spatial controls (spatial distance-weighed lags of pollution and GDP in a 100km and 200km radius around each regions centroid). Allowing for spatial autocorrelation by clustering standard errors on NUTS2 and NUTS1 regions has no bearing on the statistical significance of the result.

Weather controls. As emphasized above, our instrumental variables satisfy the exclusion restriction *conditionally*. That is, conditional on ground-level weather, both thermal inversions and wind direction should only affect economic outcomes via their effect on pollution. Because both of these variables are likely correlated with weather, which can itself impact economic outcomes, it is important to carefully control for weather. We do this in the main results using a flexible approach to controlling for temperature, precipitation, and wind speed, including other weather variables using second-degree polynomials, and interacting temperature bins with relative humidity. In the appendix, we show that our results are invariant to adopting an even more flexible approach to including the effect of temperature (conditioning on 70 temperature bins, rather than 20 as in the baseline, and interacting those 70 temperature bins with humidity and squared humidity). Moreover, we estimate the model without conditioning on any weather covariates and continue to obtain almost the same results as in the baseline specification, suggesting that our choice of weather controls is unlikely to substantially affect the results.

Alternative instruments. In our main specification, we adopt (by necessity) particular definitions for the wind direction and thermal inversion variables. Specifically, we use the number of days of thermal inversions at the lowest atmospheric level (in both winter and summer) and the number of days with winds from the three most and least polluting directions (based on pre-sample data) as instruments (a total of four instruments). We re-run the analysis with a wide range of different definitions for both of these instrumental variables (using the three different definitions of inversions described above; using annual counts or four seasons instead of two; changing the definition of the cleanest and dirtiest directions; and including dummy variables for all wind directions). We also run a specification in which we rank order the wind directions for each NUTS3 region based on the prior 5 years of daily pollution, rather than the pre-sample period. Overall, our results are robust to using a large set of different instrumental variables. The main coefficient varies between -0.0066 and -0.0133 depending on the instrument, and all different choices of instrument deliver precisely estimated and economically large estimates of the impact of pollution on output.

Alternative air pollution data. We re-estimate the model with different air pollution data, based both on the Van Donkelaar et al. (2016) and the CAMS models (see Section 4). In both cases, we continue to find a negative impact of pollution on GDP, however, the results are somewhat smaller than the benchmark estimates. It is important to note that the CAMS results are based on data with incomplete coverage of the observation period used for other results (2003-2015). We also estimate the model using data from the ground-level pollution monitoring stations. Here we do not find a statistically significant impact of pollution on economic output. It is important to note that the EEA ground-level monitoring data set is much smaller than the other data sets, due to the incomplete coverage of Europe by ground-based monitoring stations, especially early in the sample (see the above discussion). While the point estimate is not precisely estimated using the ground-based stations, it is of similar magnitude to other estimates.

Alternative specifications. We test the sensitivity of our results to the inclusion of alternative fixed effects and control variables. We add country linear time trends, country quadratic time trends, country-year fixed effects, and NUTS3 linear time trends. The coefficient remains unchanged except in the case of country-year fixed effects which remove too much variation in pollution, resulting in large standard errors and making the coefficient statistically insignificant (though not statistically different from the baseline). The reanalysis data sets that we use are heavily interpolated, such that measurement error is an important concern (but which should be alleviated by our IV approach). As we saturate the model with fixed effects, our estimates as a result are attenuated.

Co-pollutants. We control for SO_2 in an attempt to isolate the impact of $PM_{2.5}$ from other pollutants. However, instruments are weaker in predicting two pollutants together and standard errors increase. We continue to find a large negative impact of $PM_{2.5}$ on output that is close to our baseline estimate. We do not find a statistically significant impact of SO_2 on output.

Placebo test. We estimate a placebo model where we randomly assign pollution and meteorological variables from each region to a different random region in our data set, and re-estimate the model with these ”placebo” treatment variables. The placebo regression delivers a coefficient of zero on the instrumented pollution, suggesting that our results are not driven by our empirical approach, but instead by the data.

Overall, the results are robust to multiple sensitivity tests. The marginal impact of a $1\mu g/m^3$ (about 10%) increase in $PM_{2.5}$ on GDP varies between about -0.6% and -1.3% depending on the specification, suggesting that our baseline estimate of -1.1% which stands in the middle of this range is a good approximation of the true effect. We provide additional details on the robustness checks in the Appendix.

5.3 Extensions

5.3.1 Non-linearity

We next explore potential non-linearity of the effect of $PM_{2.5}$ concentration increases on economic activity, to determine if the marginal effect of pollution on economic activity is increasing or decreasing in pollution concentrations. Understanding how marginal effects change as pollution concentration changes can be helpful in extrapolating beyond this study to other regions where pollution concentrations are higher (e.g., South-east Asia) or lower (e.g., North America) than in Europe. Existing literature that studies the health impacts of air pollution is suggestive of non-linear pollution impacts. In particular, studies typically find declining marginal impacts of air pollution on health; this phenomenon is sometimes referred to as a “supra-linear” dose-response function (Pope III et al., 2015; Arceo et al., 2016).

We seek to understand potential non-linearity in the pollution-output relationship using several complementary approaches. In column (1) of Table 6, we include both a linear pollution term as well as a quadratic term in pollution as endogenous variables in the second-stage regression. In the first stage, we use the same instrumental variables as in the baseline regression, and include squares of these instruments as well. Our instruments are not as strong in this context (first stage F-stat is below 10), but the results are suggestive of non-linearity. The linear term is negative and

Robustness check		Coefficient
Weighing		
	Unweighted	-0.0129***
	GDP-weighted	-0.0101***
Excluding extreme values		
	1%	-0.0123***
	5%	-0.0174***
Spatial controls		
	Spatial lag (100km)	-0.005***
	Spatial lag (200km)	-0.009***
	NUTS2 clustered s.e.	-0.0116***
	NUTS1 clustered s.e.	-0.0116***
Weather controls		
	20 temp. bins	-0.0134***
	70 temp. bins	-0.0132***
	70 temp. bins + humidity interaction	-0.0106***
	No weather control	-0.0082***
Instrument choice		
	Inversions any (annual)	-0.0096***
	Inversions any (two seasons)	-0.0082***
	Inversions any (four seasons)	-0.0095***
	Inversions low (annual)	-0.0117***
	Inversions low (four seasons)	-0.0110***
	Inversions surface (annual)	-0.0111***
	Inversions surface (two seasons)	-0.0113***
	Inversions surface (four seasons)	-0.0116***
	Wind direction (dirtiest three directions)	-0.0133***
	Wind direction (cleanest three directions)	-0.0105***
	Winds (cleanest and dirtiest two directions)	-0.0098***
	Winds (all directions)	-0.0101***
	Winds (rolling all directions)	-0.0066***
Database choice		
	Van Donkelaar et al	-0.008***
	CAMS	-0.007***
	EEA monitoring data	-0.008
Specification		
	Linear country-trends	-0.0112***
	Quadratic country-trends	-0.0118***
	Linear NUTS3-trends	-0.0110***
	Country-year effects	-0.0041
Control for co-pollutants		
	Control for SO_2	-0.0119 ***
Placebo		
	Placebo	0.002

Table 5: Results of selected robustness checks.

the squared term is positive, pointing to a declining marginal damage from pollution, consistent with the health literature.

In column (2), we take an alternative approach to understanding non-linearity in the pollution-output relationship. Here, the key endogenous variable in our second-stage regression is the interaction of first-differenced pollution with the level of pollution. We instrument for this variable with both the baseline (first-differenced) set of instruments and interaction terms between first-differenced instruments and the corresponding instrument in level terms. We find results that are qualitatively consistent with those in column (1), such that the marginal damage from pollution appears to decline with increases in pollution.

In column (3), we explore potential non-linearity in a different way, by creating dummy variables that indicate whether pollution in each NUTS3 region is above or below the region-specific median. We interact this dummy variable with pollution in the second-stage equation. In the first-stage equation, we interact our instruments with the indicator variable as well. Again, we find evidence of declining marginal damage from pollution using this approach, with the marginal damage in above-median pollution years somewhat below the marginal damage in below-median pollution years.

Our non-linear results should be interpreted cautiously, because our instruments are less strong in some cases. However, the finding that marginal damage falls as pollution increases is consistent across our different approaches to understanding non-linearity, and is consistent with the health literature. Moreover, a declining marginal damage of pollution can help reconcile the finding that pollution causes significant damages in relatively low-pollution regions of the world with similar results in high-pollution regions.

5.3.2 Heterogeneity

Finally, we go beyond the average effect across European regions and explore the possibility of heterogeneous effects across sectors, income levels, and population density.

We begin by examining heterogeneous impacts of pollution on sector output. Figure 7 breaks the results of our estimates down by sector. In this figure, we report the results from a series of separate regressions, in which we use the gross value added of each economic sector that is reported in Eurostat as a left hand side variable in our second-stage equation. Each point reflects the change in sector gross value added (in log points) resulting from a $1\mu g/m^3$ (roughly 10%) increase in $PM_{2.5}$ concentrations. For most sectors, the estimated coefficient is similar to the economy-wide value of about -0.011 (see Table 3; this mean estimate is given by the light red line). That is, for most sectors, output is estimated to fall by about 1.1% due to a $1\mu g/m^3$ increase (10%) in pollution. There is one main outlier: we find a much larger impact of increases in air pollution on the agricultural sector than on other sectors. For this sector, a $1\mu g/m^3$ increase in $PM_{2.5}$ concentration

Table 6: Nonlinearities in $PM_{2.5}$'s effect on GDP.

	(1) Squared	(2) Interaction with level	(3) Below/above region-median
$PM_{2.5}$	-0.0457 ***	-0.0448 ***	
	(0.0108)	(0.0121)	
$PM_{2.5}^2$	0.0017 ***		
	(0.0005)		
$PM_{2.5} \times PM_{2.5}$ (level)		0.0034 ***	
		(0.0012)	
$PM_{2.5} \times$ 1(Below median)			-0.0128 ***
			(0.0026)
$PM_{2.5} \times$ 1(Above median)			-0.0100 ***
			(0.0024)
Observations	17099	17099	17099
Weak id. stat.	4.196	2.433	54.12
Hansen J stat. p-value	0.00000580	0.00000460	0.0131

Notes:* $p<0.1$,** $p<0.05$,*** $p<0.01$. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Each regression uses both wind and thermal inversions as instruments (same instruments as main regression). Column (1) includes quadratics in each instrument. Column (2) interacts first-differenced instruments with the corresponding instrument in level terms. Column (3) interacts each instrument with an indicator variable denoting whether pollution is above or below region-median in a given year. Regression coefficients weighted by each region's population.

leads to a roughly 4% reduction in sector gross value added. A number of studies have shown the sensitivity of the agriculture sector to high levels of pollution concentration, so this result can be rationalized based on the extant literature (e.g. Wahid et al., 1995; Agrawal et al., 2003; Proctor et al., 2018). Considering the framework in Section 2, it is likely that there are direct effects of pollution on output in this sector, and that the outdoor work environment exposes workers to additional pollution. In particular, there is evidence that crop output falls with increasing concentration of atmospheric $PM_{2.5}$, which scatters incoming sunlight and affects the ratio of direct to diffuse radiation. In addition, there is evidence that poor air quality causes reductions in agricultural worker productivity (Graff Zivin and Neidell, 2012), which would also affect gross value added of the sector. More generally, differences in sector impacts can be rationalized in the theoretical framework to the extent that different sectors use different combinations of inputs, are differently exposed to pollution, and differently suffer direct effects of pollution on output.

We next explore heterogeneity in results by income, population density, and by economic structure. To this end we run separate regressions on sub-groups of the data, to reveal the heterogeneous impact of air pollution for different categories of regions as described in Appendix B.2. It is important to recognize at the outset that regions in our sample differ in many ways, so that—while we are estimating a causal relationship of pollution on output—we cannot estimate causally how this relationship changes with heterogeneity, since there may be multiple dimensions of heterogeneity that co-vary.

Results differentiating regions by level of income show a weak "inverted-U" relationship, with the largest marginal effects of pollution evident in the lowest- and highest-income regions, and smaller marginal impacts of pollution in medium-income regions (Figure 16 in Appendix B.2). As described above, the "inverted-U" pattern is hardly causal, since there are many factors that are correlated with income and vary between regions. Thus the heterogeneity in causal effects of pollution across different regions represent associations rather than causal effects of heterogeneity on damages from pollution.

We next divide the sample according to population concentration into urban, rural, and "intermediate" regions (see Table 11 in Appendix B.2). The classification is adopted from the OECD and is based on population density and proximity to urban centres.[24] Impacts of air pollution are unsurprisingly concentrated in urban and "intermediate" regions where pollution concentrations are typically higher than elsewhere.

Finally, we divide the sample into industry-intensive regions and service-intensive regions, based on the sectoral contributions to total value added. We separately estimate the impact of pollution on economic output in industry- and service-intensive regions (Table 12 in Appendix

[24]See: `https://www.oecd.org/regional/regional-statistics/OECD_regional_typology_Nov2012.pdf`.

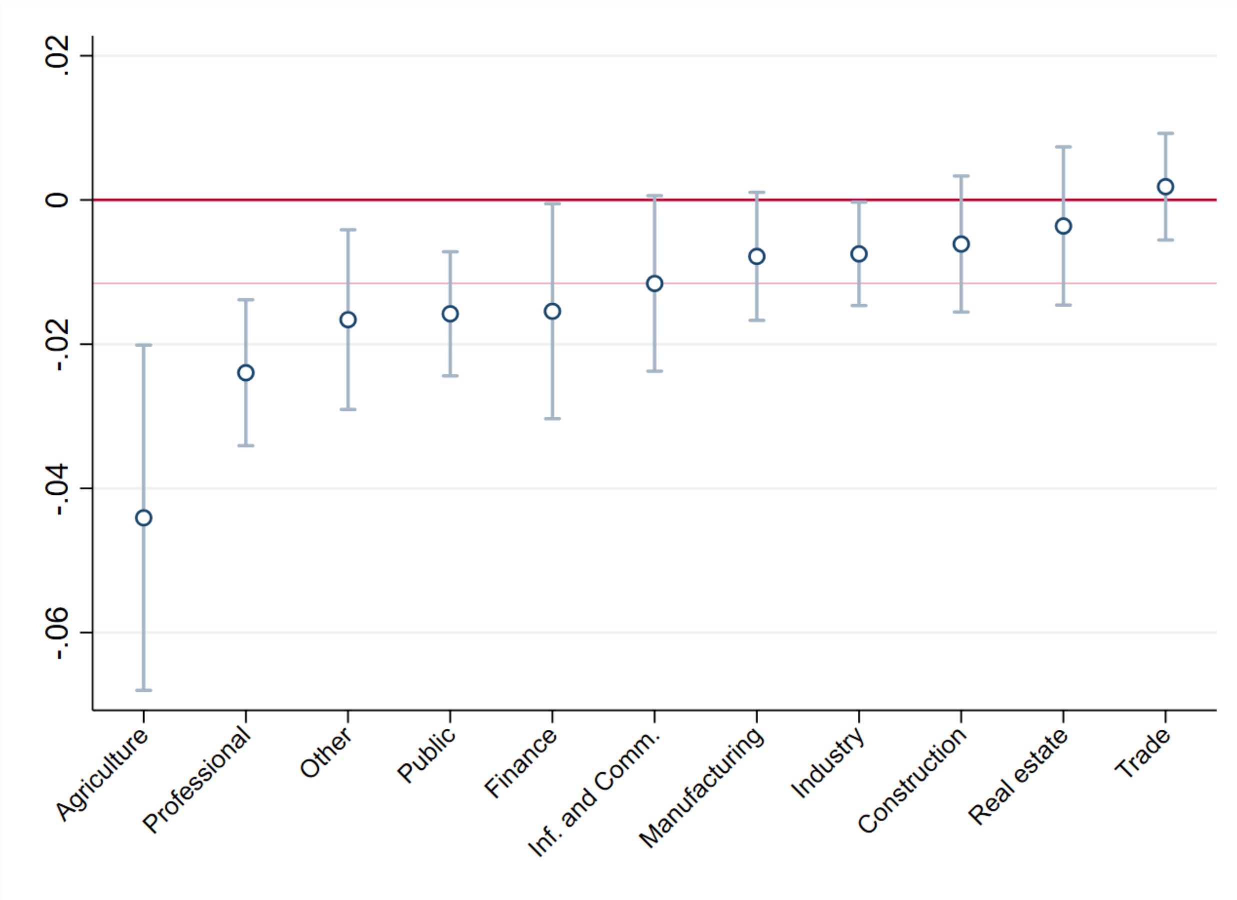

Figure 7: The impact of $PM_{2.5}$ by sector.
This chart shows the estimated coefficients and 95% confidence intervals for regressions of sector gross value added on pollution using the two stage least squares model described in the text. To generate the figure, we run separate regressions on the gross value added of each sector. The dependent variables are the log real gross value added of the given industry (based on NACE rev. 2). All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population. The line reflects the average impact across the economy estimated in Table 3. In each case, we use summer and winter thermal inversions, and clean and dirty winds as instrumental variables, as defined in the text.

B.2). We find larger impacts in service-intensive regions. This could be due to higher labour-share in these industries, since labour is more likely to be affected by pollution than capital.

6 Discussion and policy implications

6.1 Overall magnitude of finding

The most striking feature of our results is the magnitude of the effects we uncover. Our baseline (GDP-weighted) estimates show that a 1μg/m^3 or 10% decrease in $PM_{2.5}$ concentration would increase Europe's GDP by about 1%. Given that the European Union's GDP is about €15 trillion (in 2017), this translates into a short-run increase of €150 billion. This a very large number - roughly the size of a small EU Member Country such as Hungary or Slovakia. On a per capita basis, this represents around €300 per inhabitant per year. Another way to put things into perspective is to note that pollution decreased by 0.7% per year on average across Europe between 2000 and 2015, so in a typical year, reduction in pollution boosts regional GDP by 0.07%. As a matter of comparison, regional GDP (at constant prices) grew by 0.7% per year on average over the same period, so reductions in air pollution explain about 10% of GDP growth, according to our results.

6.2 Comparison with existing studies

There are few studies that undertake to estimate the impact of air pollution on the overall level of market economic activity, but our results are remarkably consistent with some of those reported in this emerging literature. Fu et al. (2017) estimate that just a 2% increase in $PM_{2.5}$ concentration causes labour productivity to decrease by 1.1% in Chinese manufacturing plants. Our results are much *smaller* in magnitude, suggesting that air pollution matters even at much lower concentration levels than those observed in China but that the effect is clearly non-linear. Borgschulte et al. (2020) focus on pollution peaks in the U.S. caused by forest fires. They estimate that spending one day in a smoke plume causes a reduction in income of 10% across all workers. They estimate that smoke increases $PM_{2.5}$ concentration by 4μg/m^3. Thus, a 1μg/m^3(about 10%) increase in pollution causes a 2.5% reduction in income, which is much *higher* than our baseline estimates. Finally, Isen et al. (2017) estimate that a 10% increase in total particulate matter exposure during childhood reduces income 30 years later by about 1%. There is clearly a different mechanism at play in this latter study compared to the others described above and to ours, but the consistency in the effect size in comparison to our study is notable.

6.3 Implications for cost-benefit analyses of pollution control policies

These findings can inform ex-ante and ex-post cost-benefit evaluations of air pollution reduction policies. They suggest that the direct economic benefits from air pollution control policies might be much greater than previously thought, and are also much larger than abatement costs.

6.3.1 Comparison with market benefits attributed to pollution reductions

It is useful to compare the market benefits from air pollution reductions uncovered in this study with market benefits estimated in existing cost-benefit analyses of air pollution control policies. A first example is a recent assessment carried out by the European Commission when it proposed a new Directive to further reduce emission of certain atmospheric pollutants in Europe by the year 2025 (European Commission, 2013). The scenarios analysed focus on reductions in $PM_{2.5}$ emissions by 17% to 45%. The estimated market benefits from reduced $PM_{2.5}$ emissions included benefits due to reductions in lost working days, damage to the built environment, crop value losses, and healthcare costs (notably, possible changes in on-the-job productivity are not included in these calculations). The direct market benefits from reducing $PM_{2.5}$ emissions by 17% were estimated to be €1 billion annually, and around €2 billion for a 25% reduction (see details in Table 25 in Appendix C). Therefore, the direct market benefits from a 10% reduction in emissions as estimated by the European Commission are less than €1 billion. In contrast, our empirical estimates conclude that the market benefits from a 10% reduction in pollution are over €100 billion — or two orders of magnitude larger. While we cannot identify the source of the disparity precisely, it seems likely that omissions of on-the-job productivity benefits from reduced pollution in the EU study is the source of at least part of the difference.

Another example is the assessment of the costs and benefits of the U.S. Clean Air Act Amendments (CAAA) conducted by the U.S. Environment Protection Agency (EPA) (US EPA, 2011). Here, the market benefits included minor restricted activity days, work loss days, reduced outdoor worker productivity, and agricultural and forest productivity, and the US EPA estimates that the combined benefits amount to $20.5 billion annually (see details in Table 26 in Appendix C). Since the US CAAA led to a reduction in $PM_{2.5}$ emissions by 11% in 2010 compared to a scenario without CAAA (and -17% in 2020), the results in this paper suggest that the market benefits would in fact be in the order of $100 billion annually (assuming similar causal effects of pollution on output in the US).

Overall, the results from this study suggest that prior estimates of the market benefits of pollution abatement have been substantially underestimated in prior assessments.

6.3.2 Comparison with non-market benefits attributed to pollution reductions

In typical cost-benefit analyses of public policies to reduce air pollution, benefits are dominated by non-market benefits, and in particular by estimates of the value of reduced mortality. For example, in the European Commission cost benefit study of air pollution reductions, non-market benefits (in particular reduced mortality) are estimated to be 93-98% of total benefits from pollution reductions (European Commission, 2013). In the US EPA study of the Clean Air Act Ammendments, non-market benefits are estimated to be around 99% of total benefits from pollution reduction (US EPA, 2011).[25]

In contrast, this study finds that market benefits of pollution reductions are of comparable magnitude to non-market benefits. The World Health Organisation estimates that outdoor air pollution is responsible for over 400,000 annual deaths in the European Union. Let us make the reasonable assumption that reducing $PM_{2.5}$ concentration by 10% would reduce the number of deaths by 10%. [26] Using the base Value of Statistical Life of US$3 million for OECD countries (OECD, 2016), the monetized benefits from avoided mortality caused by a 10% decrease in $PM_{2.5}$ concentration amount to US$ 120 billion annually, or €110 billion, a figure which is comparable in magnitude to the market economic benefit found in this paper. We conclude from this back of the envelope calculation that including the direct economic benefits of air pollution control into cost-benefit analyses of policies would lead to increase the expected benefits from policy action usually projected by around 50%.

6.3.3 Comparison with abatement costs

How do our estimates of the market benefits of pollution reduction compare to the marginal abatement costs of decreasing pollution? The European Commission has conducted an impact assessment of a policy package aimed at further reducing emission of pollutants in Europe by the year 2025 (European Commission, 2013). Unfortunately, the scenarios analysed focus on emission reductions rather than decrease in concentration, and it is not easy to translate emission reductions into concentration. One cannot expect a perfectly linear relationship between the reductions in emissions of primary $PM_{2.5}$ and the reductions in ambient air concentrations, because in addition to primary emissions of particles, $PM_{2.5}$ can also be formed from the chemical reactions of gases such as SO_2 and NO_X, and because wind can transport particles over long distances. However, between 2006 and 2014, primary $PM_{2.5}$ emissions decreased by 17% in the EU28 while in the same period, PM2.5 concentrations declined by 20% on average (indicating a small reduction in

[25]See Tables 25 and 26 Appendix C for details.

[26]The European Commission's 2013 impact assessment finds such a close to linear relationship. For example, Scenario A leads to a 36% reduction in $PM_{2.5}$emissions and a 42% reduction in premature deaths (European Commission, 2013).

secondary PM also). Therefore, as a first approximation, it is not unreasonable to assume a linear relationship between emissions and concentration, especially for a large region such as Europe.

Table 27 in Appendix C reports the results of the European Commission impact assessment analysis. This cost-benefit study suggests the marginal cost of mitigating $PM_{2.5}$ emissions by about 17% would be €221 million annually. Similarly, the marginal cost of a 25% reduction would be about €1.2 billion. Thus it is reasonable to assume, based on this study, that the cost of a 10% reduction in emissions would be less than €1 billion. In contrast, our estimates suggest that the market benefit of a 10% reduction in emissions would be roughly two orders of magnitude larger.

The US EPA estimates are larger with the annual abatement costs associated with the CAAA amounting to $65 billion annually (Table 28 Appendix C), but these numbers include abatement of many pollutants other than $PM_{2.5}$ (e.g. NOX, CO, SO2, PM10) and even then these numbers are around twice as small as our estimated direct market benefits.

We conclude from this analysis that significant reductions in air pollution would easily pass a cost-benefit test, even ignoring their large benefits in terms of avoided mortality. Therefore, more stringent air quality regulations could be warranted based solely on economic grounds.

7 Conclusions

In this paper we have combined data on sub-national GDP and other economic outcomes, air pollution, and weather conditions to estimate the causal effect of air pollution on economic activity in Europe. We use data on local GDP for 1,342 NUTS-3 regions over 16 years (2000-2015) that we match with air pollution and weather data . To circumvent the problem of reverse causality, our identification strategy relies on quasi-random variation in thermal inversions and wind direction as instruments for air pollution. Thermal inversions are upper-atmosphere phenomena that are exogenous to economic activity but trap pollution over the earth surface. Wind blowing in different directions causes pollution to be imported or exported from neighbouring regions. With this, we provide the first causal evidence on the impact of air pollution on economic activity in Europe.

We find that increases in air pollution cause substantial reductions in economic activity. Our baseline model shows that reducing the average concentration of fine particulate pollution by 10% would cause economic activity to increase by 1.1%, or about €150 billion—roughly the size of a small EU Member Country such as Hungary or Slovakia. On a per capita basis, this amounts to around €300 per person per year. We disaggregate the effect into population and productivity and find that 95% of the impact comes from reduced productivity and 5% to reduced population. We explore the heterogeneity of this impact across economic sectors and find large impacts in agriculture—consistent with the fact that workers in this sector operate mostly outdoors and that crop production can be directly impacted by air pollution. However, we also find negative impacts of air pollution on high-skill sectors where workers are indoors, such as finance and manufacturing.

We view these results as of substantial importance to evaluations of the benefits of reducing air pollution. Air pollution regulations are typically scrutinized by comparing costs to benefits. In most evaluations, the benefits of air pollution regulations are dominated by impacts on mortality. Economic benefits—such as absenteeism at work—are normally considered to be of second order importance in these evaluations. Our results suggest instead that the economic impacts of air pollution are substantial and of similar magnitude to benefits from reduced mortality. In other words, air pollution might be much more costly than normally believed, and much stronger air quality regulations could be warranted.

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham, "Management and shocks to worker productivity," Technical Report, National Bureau of Economic Research 2019.

Agrawal, M., B. Singh, M. Rajput, F. Marshall, and J.N.B Bell, "Effect of air pollution on peri-urban agriculture: a case study," *Environmental Pollution*, 2003, *126* (3), 323 – 329.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2008.

Aragon, Fernando M, Juan Jose Miranda, and Paulina Oliva, "Particulate matter and labor supply: the role of caregiving and non-linearities," *Journal of Environmental Economics and Management*, 2017.

Arceo, Eva, Rema Hanna, and Paulina Oliva, "Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City," *The Economic Journal*, 2016, *126* (591), 257–280.

Archsmith, James, Anthony Heyes, and Soodeh Saberian, "Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation," *Journal of the Association of Environmental and Resource Economists*, 2018, *5* (4), 827–863.

Auffhammer, Maximilian, "Quantifying economic damages from climate change," *Journal of Economic Perspectives*, 2018, *32* (4), 33–52.

Avnery, Shiri, Denise L Mauzerall, Junfeng Liu, and Larry W Horowitz, "Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses and economic damage," *Atmospheric Environment*, 2011, *45* (13), 2284–2296.

Bondy, Malvina, Sefi Roth, and Lutz Sager, "Crime is in the air: The contemporaneous relationship between air pollution and crime," *Journal of the Association of Environmental and Resource Economists*, 2020, *7* (3), 555–585.

Borgschulte, Mark, David Molitor, and Eric Zou, "Air pollution and the labor market: Evidence from wildfire smoke," Technical Report, Technical report. Working Paper 2020.

Botta, Enrico and Tomasz Koźluk, "Measuring environmental policy stringency in OECD countries," 2014.

Buchard, V, AM da Silva, CA Randles, P Colarco, R Ferrare, J Hair, C Hostetler, J Tackett, and D Winker, "Evaluation of the surface PM 2.5 in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States," *Atmospheric Environment*, 2016, *125*, 100–111.

_ , CA Randles, AM da Silva, A Darmenov, PR Colarco, R Govindaraju, R Ferrare, J Hair, AJ Beyersdorf, LD Ziemba et al., "The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies," *Journal of Climate*, 2017, *30* (17), 6851–6872.

Burke, Marshall, Solomon M Hsiang, and Edward Miguel, "Global non-linear effect of temperature on economic production," *Nature*, 2015, *527* (7577), 235–239.

Calderón-Garcidueñas, Lilian, Ricardo Torres-Jardón, Randy J Kulesza, Su-Bin Park, and Amedeo DAngiulli, "Air pollution and detrimental effects on childrens brain. The need for a multidisciplinary approach to the issue complexity and challenges," *Frontiers in human neuroscience*, 2014, *8*, 613.

Chameides, William L, H Yu, SC Liu, M Bergin, X Zhou, L Mearns, G Wang, CS Kiang, RD Saylor, C Luo et al., "Case study of the effects of atmospheric aerosols and regional haze on agriculture: an opportunity to enhance crop yields in China through emission controls?," *Proceedings of the National Academy of Sciences*, 1999, *96* (24), 13626–13633.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell, "Particulate Pollution and the Productivity of Pear Packers," *American Economic Journal: Economic Policy*, 2016, *8* (3), 141–69.

Chang, Tom Y, Joshua Graff Zivin, Tal Gross, and Matthew Neidell, "The effect of pollution on worker productivity: evidence from call center workers in China," *American Economic Journal: Applied Economics*, 2019, *11* (1), 151–72.

_ , Wei Huang, and Yongxiang Wang, "Something in the air: Pollution and the demand for health insurance," *The Review of Economic Studies*, 2018, *85* (3), 1609–1634.

Chay, Kenneth Y and Michael Greenstone, "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession," *The quarterly journal of economics*, 2003, *118* (3), 1121–1167.

Chen, Shuai, Paulina Oliva, and Peng Zhang, "The Effect of Air Pollution on Migration: Evidence from China," Technical Report 24036, National Bureau of Economic Research 2017.

Cohen, Aaron J, Michael Brauer, Richard Burnett, H Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan, Bert Brunekreef, Lalit Dandona, Rakhi Dandona et al., "Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015," *The Lancet*, 2017, *389* (10082), 1907–1918.

Currie, Janet and Matthew Neidell, "Air pollution and infant health: what can we learn from California's recent experience?," *The Quarterly Journal of Economics*, 2005, *120* (3), 1003–1030.

_ , Eric A Hanushek, E Megan Kahn, Matthew Neidell, and Steven G Rivkin, "Does pollution increase school absences?," *The Review of Economics and Statistics*, 2009, *91* (4), 682–694.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken, "Temperature and income: reconciling new cross-sectional and panel estimates," *American Economic Review*, 2009, *99* (2), 198–204.

_ , _ , and _ , "Temperature shocks and economic growth: Evidence from the last half century," *American Economic Journal: Macroeconomics*, 2012, *4* (3), 66–95.

Deryugina, Tatyana and Solomon Hsiang, "The Marginal Product of Climate," Technical Report 24072, National Bureau of Economic Research 2017.

_ and Solomon M Hsiang, "Does the environment still matter? Daily temperature and income in the United States," Technical Report, National Bureau of Economic Research 2014.

_ , Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif, "The mortality and medical costs of air pollution: Evidence from changes in wind direction," *American Economic Review*, 2019, *109* (12), 4178–4219.

Dingenen, Rita Van, Frank J Dentener, Frank Raes, Maarten C Krol, Lisa Emberson, and Janusz Cofala, "The global impact of ozone on agricultural crop yields under current and future air quality legislation," *Atmospheric Environment*, 2009, *43* (3), 604–618.

Dockery, Douglas W, C Arden Pope, Xiping Xu, John D Spengler, James H Ware, Martha E Fay, Benjamin G Ferris Jr, and Frank E Speizer, "An association between air pollution and mortality in six US cities," *New England journal of medicine*, 1993, *329* (24), 1753–1759.

Donkelaar, Aaron Van, Randall V Martin, Michael Brauer, N Christina Hsu, Ralph A Kahn, Robert C Levy, Alexei Lyapustin, Andrew M Sayer, and David M Winker, "Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors," *Environmental science & technology*, 2016, *50* (7), 3762–3772.

Du, Yixing, Xiaohan Xu, Ming Chu, Yan Guo, and Junhong Wang, "Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence," *Journal of thoracic disease*, 2016, *8* (1), E8.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth, "The long-run economic consequences of high-stakes examinations: evidence from transitory variation in pollution," *American Economic Journal: Applied Economics*, 2016, *8* (4), 36–65.

European Commission, "DIRECTIVE 2008/50/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 May 2008 on ambient air quality and cleaner air for Europe," Technical Report 2008.

_ , "Executive summary of the impact assessment," Technical Report SWD(2013) 532 final 2013.

Fu, Shihe, V Brian Viard, and Peng Zhang, "Air quality and manufacturing firm productivity: comprehensive evidence from China," 2017.

Graff Zivin, Joshua and Matthew Neidell, "The impact of pollution on worker productivity," *The American Economic Review*, 2012, *102* (7), 3652–3673.

_ **and** _ , "Environment, health, and human capital," *Journal of Economic Literature*, 2013, *51* (3), 689–730.

_ **and** _ , "Air pollution's hidden impacts," *Science*, 2018, *359* (6371), 39–40.

Gramsch, E., D. Cceres, P. Oyola, F. Reyes, Y. Vsquez, M.A. Rubio, and G. Snchez, "Influence of surface and subsidence thermal inversion on PM2.5 and black carbon concentration," *Atmospheric Environment*, 2014, *98*, 290 – 298.

Hanna, Rema and Paulina Oliva, "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City," *Journal of Public Economics*, 2015, *122*, 68–79.

Hansen, Anett C and Harald K Selte, "Air pollution and sick-leaves," *Environmental and Resource Economics*, 2000, *16* (1), 31–50.

Haylock, MR, N Hofstra, AMG Klein Tank, EJ Klok, PD Jones, and M New, "A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006," *Journal of Geophysical Research: Atmospheres*, 2008, *113* (D20).

He, Jiaxiu, Haoming Liu, and Alberto Salvo, "Severe air pollution and labor productivity: Evidence from industrial towns in China," *American Economic Journal: Applied Economics*, 2019, *11* (1), 173–201.

Herrnstadt, Evan and Erich Muehlegger, "Air pollution and criminal activity: Evidence from Chicago microdata," Technical Report 21787, National Bureau of Economic Research 2015.

Heyes, Anthony, Matthew Neidell, and Soodeh Saberian, "The Effect of Air Pollution on Investor Behavior: Evidence from the S&P 500," Technical Report 22753, National Bureau of Economic Research 2016.

Hicks, Daniel, Patrick Marsh, and Paulina Oliva, "Air pollution and procyclical mortality: Causal evidence from thermal inversions," Technical Report, working paper. Jacobson, M.(2002). Atmospheric Pollution. History, Science, and Regulation, 1st ed. Cambridge, UK: Cambridge University Press 2016.

Holub, Felix, Laura Hospido, and Ulrich J Wagner, "Air pollution and labor supply: Evidence from social security data," 2016.

III, C Arden Pope, Maureen Cropper, Jay Coggins, and Aaron Cohen, "Health benefits of air pollution abatement policy: Role of the shape of the concentration–response function," *Journal of the Air & Waste Management Association*, 2015, *65* (5), 516–522.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker, "Every Breath You TakeEvery Dollar Youll Make: The Long-Term Consequences of the Clean Air Act of 1970," *Journal of Political Economy*, 2017, *125* (3), 848–902.

Jayachandran, Seema, "Air quality and early-life mortality evidence from Indonesias wildfires," *Journal of Human resources*, 2009, *44* (4), 916–954.

Li, Xiaoyuan, Fabian Wagner, Wei Peng, Junnan Yang, and Denise L. Mauzerall, "Reduction of solar photovoltaic resources due to air pollution in China," *Proceedings of the National Academy of Sciences*, 2017, *114* (45), 11867–11872.

OECD, *The Economic Consequences of Outdoor Air Pollution* 2016.

Pope, C Arden III, Richard T Burnett, Michael J Thun, Eugenia E Calle, Daniel Krewski, Kazuhiko Ito, and George D Thurston, "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution," *Jama*, 2002, *287* (9), 1132–1141.

Proctor, Jonathan, Solomon Hsiang, Jennifer Burney, Marshall Burke, and Wolfram Schlenker, "Estimating global agricultural effects of geoengineering using volcanic eruptions," *Nature*, 2018, *560* (7719), 480.

Ranft, Ulrich, Tamara Schikowski, Dorothee Sugiri, Jean Krutmann, and Ursula Krämer, "Long-term exposure to traffic-related particulate matter impairs cognitive function in the elderly," *Environmental research*, 2009, *109* (8), 1004–1011.

Ransom, Michael R and C Arden Pope III, "Elementary school absences and PM10 pollution in Utah Valley," *Environmental research*, 1992, *58* (1-2), 204–219.

Rémy, S, A Benedetti, A Bozzo, T Haiden, L Jones, M Razinger, J Flemming, RJ Engelen, VH Peuch, and JN Thepaut, "Feedbacks of dust and boundary layer meteorology during a dust storm in the eastern Mediterranean," *Atmospheric Chemistry and Physics*, 2015, *15* (22), 12909.

Sager, Lutz, "Estimating the effect of air pollution on road safety using atmospheric temperature inversions," *Journal of Environmental Economics and Management*, 2019, *98*, 102250.

Schlenker, Wolfram and W Reed Walker, "Airports, air pollution, and contemporaneous health," *The Review of Economic Studies*, 2015, *83* (2), 768–809.

Schulze, E-D, "Air pollution and forest decline in a spruce (Picea abies) forest," *Science*, 1989, *244* (4906), 776–783.

US EPA, "The benefits and costs of the Clean Air Act from 1990 to 2020. Final report – Rev. A," Technical Report 2011.

Wahid, A., R. Maggs, S.R.A. Shamsi, J.N.B. Bell, and M.R. Ashmore, "Air pollution and its impacts on wheat yield in the Pakistan Punjab," *Environmental Pollution*, 1995, *88* (2), 147 – 154.

Wallace, Julie and Pavlos Kanaroglou, "The effect of temperature inversions on ground-level nitrogen dioxide (NO2) and fine particulate matter (PM2.5) using temperature profiles from the Atmospheric Infrared Sounder (AIRS)," *Science of The Total Environment*, 2009, *407* (18), 5085 – 5095.

Ward, Courtney J, "It's an ill wind: The effect of fine particulate air pollution on respiratory hospitalizations," *Canadian Journal of Economics/Revue canadienne d'économique*, 2015, *48* (5), 1694–1732.

World Health Organization, "Ambient air pollution: A global assessment of exposure and burden of disease," Technical Report 2016.

For Online Publication

A Data details

A.1 Source of data

Additional details relating to the construction of all the variables used in the econometric analysis are provided in the table below.

Variable	**Construction**
Dependent variables	
GDP	Regional gross domestic product in current prices is obtained from the Eurostat data catalogue. We obtain annual data for each NUTS3 region from 2000-2015 (Eurostat table: nama_10r_3gdp). We calculate the real 2015 values using the Harmonised Index of Consumer Prices (HICP) available from Eurostat (Eurostat table: prc_hicp_aind; variable: INX_A_AVG).
GVA	Gross value added by sector is obtained from Eurostat (table: nama_10r_3gva).
Population	Population data is from Eurostat (table: demo_r_pjanaggr3).
Independent variables	
$PM_{2.5}$	We obtain daily mean $PM_{2.5}$ concentration for each grid cell covering the bounding box defined by the (longitude,latitude) coordinates (-15,35) to (35,70) from the M2T1NXAER MERRA2 reanalysis files, spanning the range January 1, 1995 to December 31, 2015. We construct a measure of $PM_{2.5}$ from the five individual species of particulate matter reported by MERRA following Buchard et al. (2016) as described in the main text. We obtain a daily measure of $PM_{2.5}$ in each NUTS3 region as the mean of all grid cells overlapping the region. We obtain an annual average concentration as the mean over all days of the year.
SO_2	Construction of SO_2 concentration follows the same procedure and with the same data source as for $PM_{2.5}$.

Temperature	We obtain daily mean temperature from the European Climate Assessment for grid cells spanning the bounding box defined by the (longitude,latitude) coordinates (-15,35) to (35,70). We obtain a daily measure of temperature in each NUTS3 region as the mean of all grid cells overlapping the region. We cut the continuous temperature into a number of temperature bins that span the range of observed temperatures. We count the number of days that mean daily temperature falls into each of these bins in each year and regions in the sample.
Precipitation	Precipitation data is derived from the same source as temperature data, and variable construction generally follows an identical procedure. We construct mean precipitation across all days of the year.
Relative humidity	We obtain relative humidity from the MERRA2 M2I3NPASM reanalysis files and follow the same procedure as described above to arrive at an annual mean relative humidity.
Surface pressure	We obtain surface pressure from the European Climate Assessment files and follow the same procedure as described above to arrive at an annual mean pressure.
Wind speed and direction	We obtain daily mean easterly and northerly wind speeds at surface from the MERRA2 M2I3NPASM reanalysis files. We obtain a daily measure of wind speed and direction in each NUTS3 region as the mean of all grid cells overlapping the region. We count the number of days that wind speed falls into bins defined by the Beaufort scale in each year and NUTS3 region. We count the number of days that wind originates from each of eight compass octants in each year and NUTS3 region.

A.2 Residual variation

Figures 8 and 9 show the identifying variation in two of the instrumental variables—thermal inversions and wind direction. Both figures show a histogram of these variables after first differencing and conditioning on year fixed effects. The figures indicate that even after removing these fixed effects, there remains a substantial amount of variation in these variables. For example, it is normal to observe 5% or 365×0.0=18 days more or less of thermal inversions than the average (i.e., the standard deviation in the share of days with inversions is approximately 5%). A similar amount of variation exists for the wind instrument. It is evident there exists considerable variation in these instruments, which we leverage for identification in our regressions. Figures 10 to 13 also report the distribution of geographic variance in the instrumental variables. The figures show that wind direction is most variable throughout central Europe, such that these regions will contribute most to the identification of the effect of the impact of pollution on GDP using this instrument. In contrast, Northern Europe and the Baltic region experience the most variation in winter inversions from year to year, and the coastal areas experience the most variation in summer inversions. These regions will contribute most to identification of the impact of pollution on GDP using the inversion instrumental.

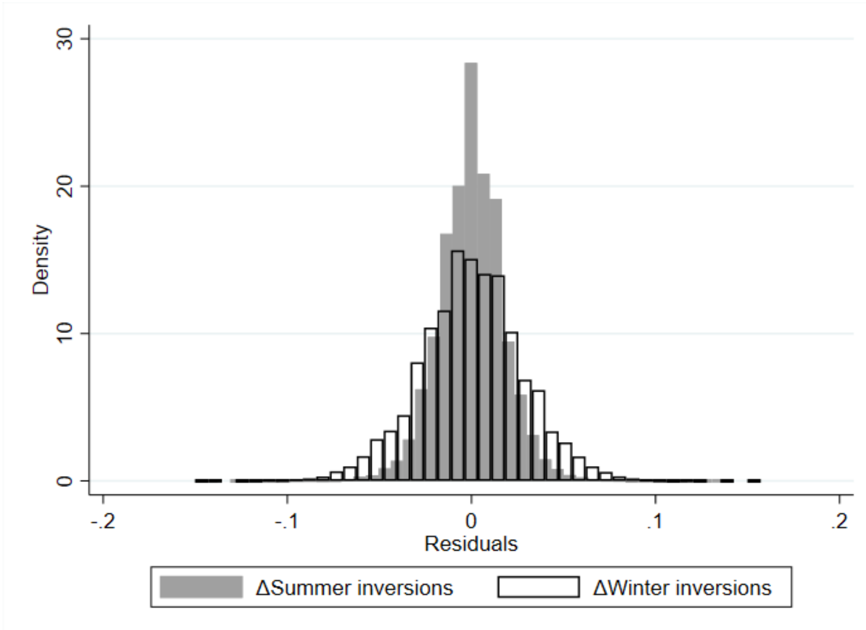

Figure 8: Histogram of inversion residuals after first differencing and conditioning on year fixed effects. The inversion variable is TI^L as described above. The variable measures the share of days in a year with thermal inversions, for both summer and winter.

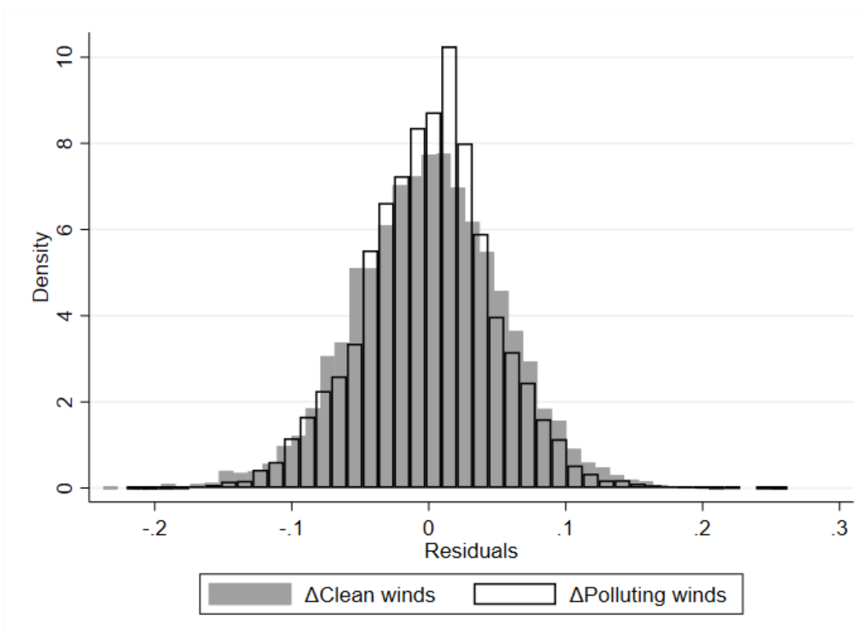

Figure 9: Histogram of wind direction residuals after first differencing and conditioning on year fixed effects. The variable measures the share of days in a year with winds originating from one of the three dirtiest or three cleanest compass octants.

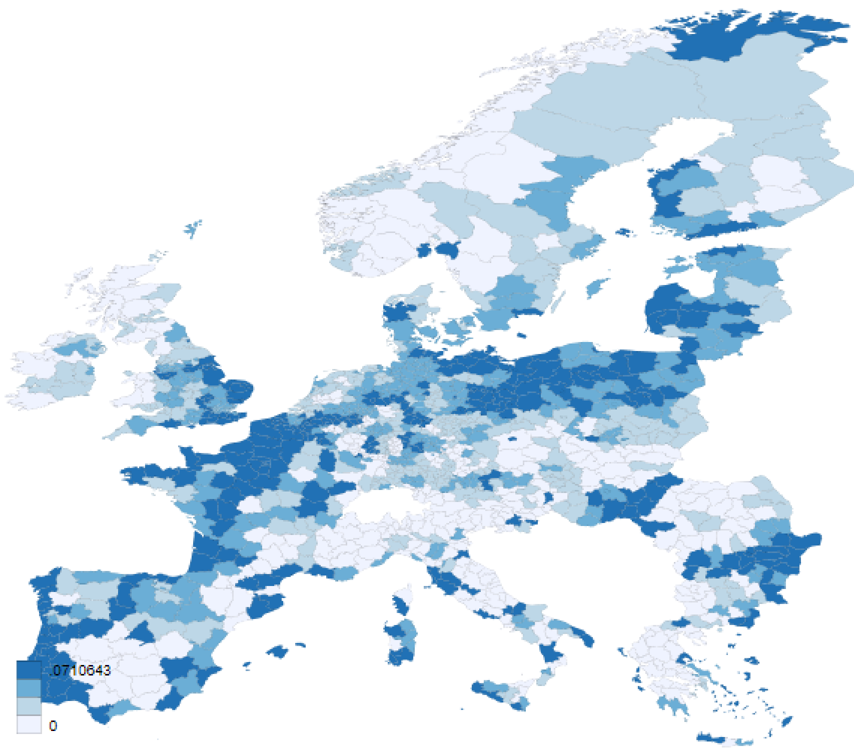

Figure 10: Geographic variation in the low level summer inversions instrument (TI^L). The figure is produced by calculating the standard deviation of the first-differenced annual inversion frequency separately in each NUTS3 region.

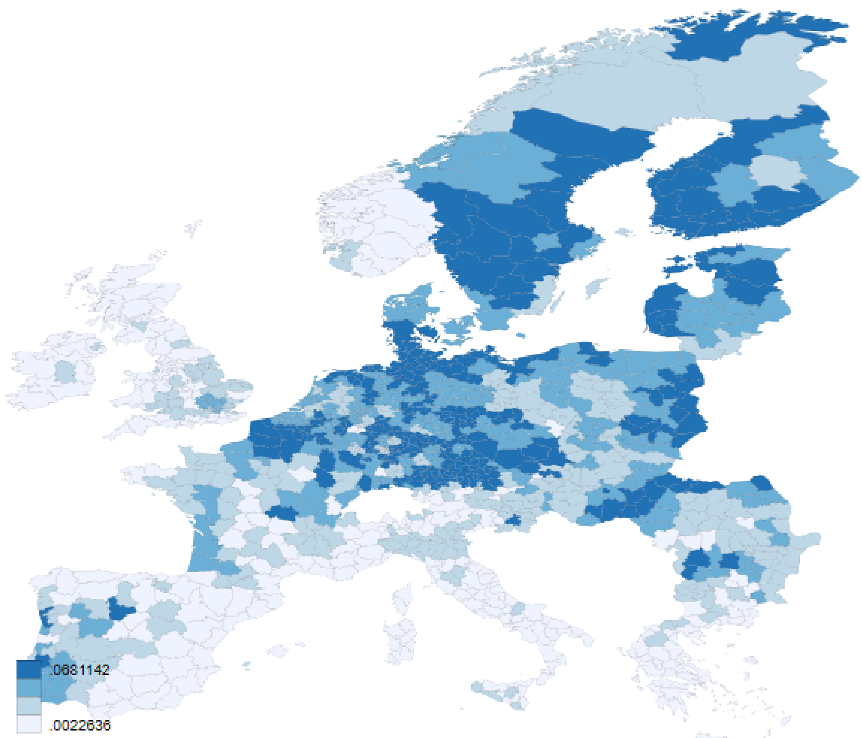

Figure 11: Geographic variation in the low level winter inversions instrument (TI^L). The figure is produced by calculating the standard deviation of the first-differenced annual inversion frequency separately in each NUTS3 region.

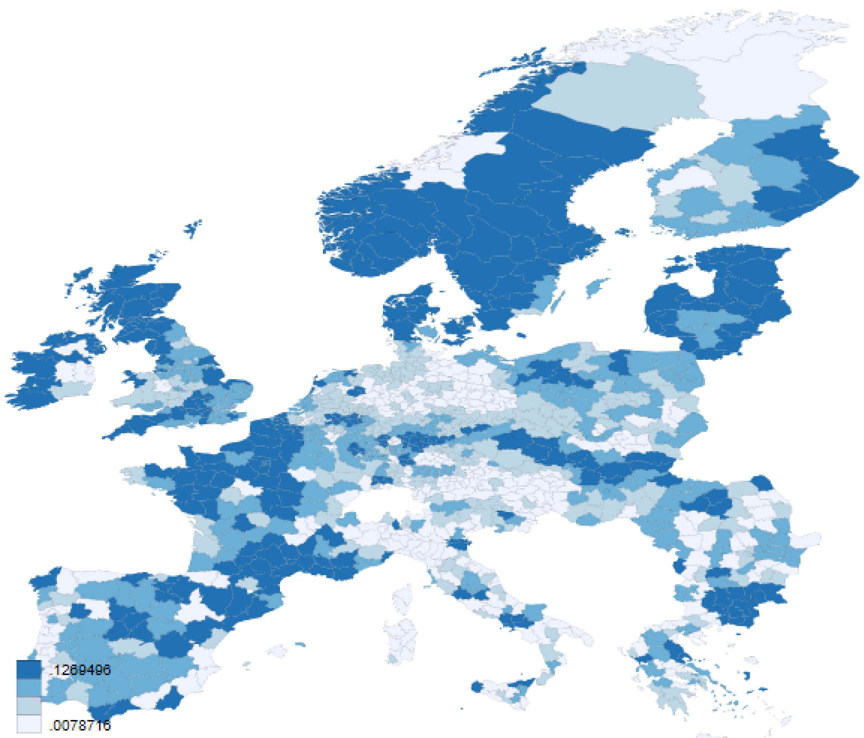

Figure 12: Geographic variation in the polluting wind instrument (wind from three dirtiest compass octants). The figure is produced by calculating the standard deviation of the first-differenced annual wind direction frequency separately in each NUTS3 region.

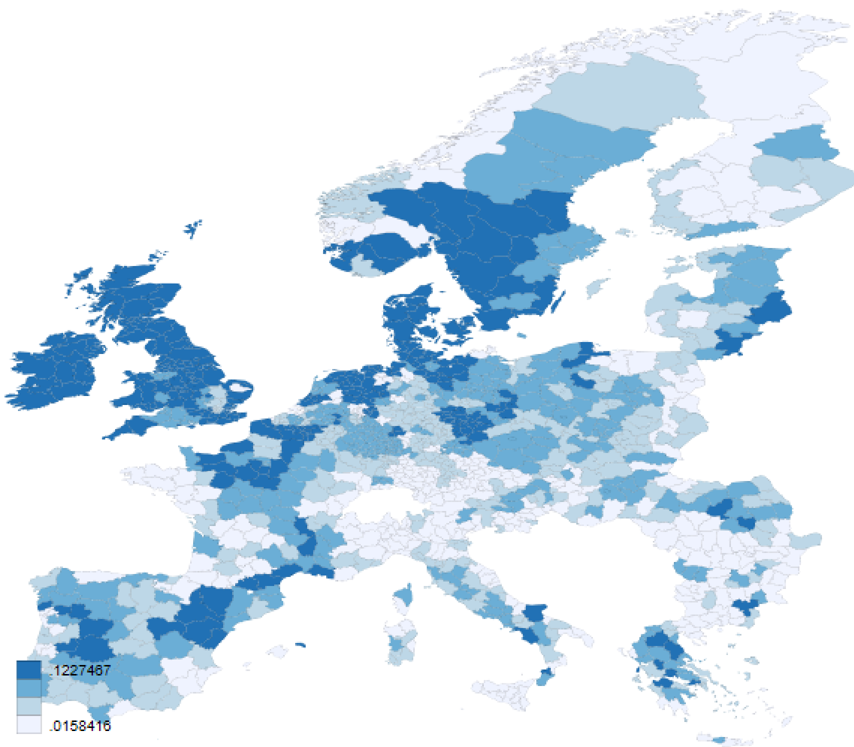

Figure 13: Geographic variation in the clean wind instrument (wind from three dirtiest compass octants). The figure is produced by calculating the standard deviation of the first-differenced annual wind direction frequency separately in each NUTS3 region.

A.3 Summary statistics

Table 8: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Real GDP	9660.491	15255.596	206.726	204152.375	18386
$PM_{2.5}$	10.870	2.172	3.284	25.059	18386
Population	373894.927	423228.554	19744	6425522	18386
Inversions (low; summer)	0.05	0.038	0	0.348	18386
Inversions (low; winter)	0.118	0.064	0	0.419	18386
Clean winds	0.428	0.114	0.074	0.773	18386
Polluting winds	0.305	0.09	0.068	0.77	18386
Surface relative humidity	0.734	0.07	0.464	0.881	18386
Pressure	1015.922	2.103	1002.326	1030.665	18386
Temperature	10.289	2.594	-1.895	19.016	18386
Precipitation	2.119	0.772	0.251	7.485	18386
Wind speed	6.185	1.223	1.896	10.16	18386

A.4 Correlation between instrumental variables

Table 9 provides a pair-wise correlation matrix for the inversion variables. Correlations are calculated after first-differencing, and removing year fixed effects, for consistency with the empirical approach.

Table 9: Correlation between main instrumental variables.

	Summer inversions (low)	Winter inversions (low)	Clean winds
Winter inversions (low)	0.0589		
Clean winds	-0.1190	-0.1122	
Polluting winds	0.1229	0.1572	-0.6528

Notes: Correlation coefficients are calculated after first-differencing and conditional on year fixed effects.

A.5 Ground-based monitoring stations

We obtain ground-based pollution monitoring data from the European Environment Agency. At the beginning of our sample period, only a handful of $PM_{2.5}$ pollution monitors were active in Europe. The network has since expanded to between 1,500 and 1,750 stations. We map pollution monitors to NUTS-3 region. The number of NUTS-3 regions with at least one active pollution monitor increased from almost none in 2000 to about 500 by 2015 (slightly over 1/3 of the total number of NUTS-3 regions in Europe). See Figure 14.

Figure 14: Number of ground-based monitoring stations (left panel) and number of NUTS-3 regions containing at least one ground-based monitoring station (right panel) from 2000-2015. Note that there is no data available for 2012 from the European ground based-monitoring network.

Figure 15 plots the relationship between the ground-based monitoring stations and the MERRA-2 $PM_{2.5}$ concentrations. Although the EEA stations—which are typically located in urban areas—measure higher levels of pollution than the MERRA-2 reanalysis—which is not population-weighted—the two measures are strongly correlated (the correlation between the two measures is about 0.56).

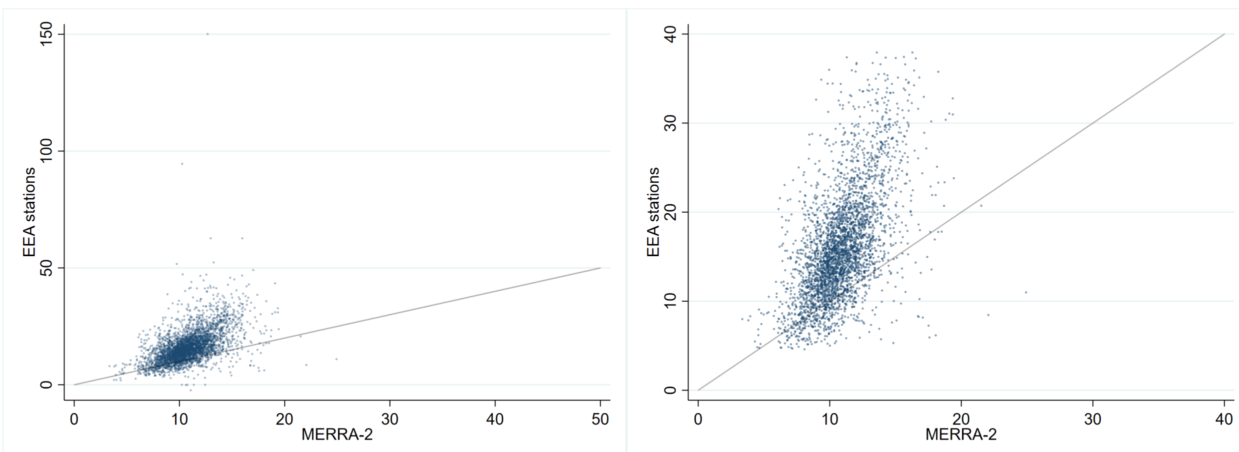

Figure 15: Correlation between MERRA-2 and ground-based monitoring station $PM_{2.5}$ concentrations. The figure on the right removes the top 1% of outliers.

B Robustness checks and additional results

B.1 Reduced-form results: The effect of wind direction and thermal inversions on economic output

This section show reduced-form results. In this case, we estimate the effect of thermal inversions and wind direction directly on economic output (conditional on weather and fixed effects). While these results are of less policy relevance, they help to build the case for the relevance of our chosen instrumental variables.

Table 10 presents the reduced-form results. In line with Table 2, column (1) uses thermal inversions as the instrument, column (2) uses wind direction, and column (3) uses both sets of instruments simultaneously. All specifications suggest that the atmospheric phenomena that we employ as instrumental variables—and which we have shown cause increases in pollution—cause negative impacts on economic activity. Specifically, increasing the share of summer inversion days by 1 percentage point (i.e., an additional $\frac{365}{100} = 3.65$ summer inversion days per year) is estimated to cause a 0.06% reduction in economic activity. Likewise, increasing the share of days in which the prevailing wind originates from the most polluting directions by 1 percentage point is estimated to cause a 0.04% reduction in economic activity.

Table 10: Reduced form: instruments' effect on the GDP.

	(1) ln(GDP)	(2) ln(GDP)	(3) ln(GDP)
Summer inversions (low)	-0.0741 ***		-0.0596 **
	(0.0258)		(0.0266)
Winter inversions (low)	0.0068		0.0013
	(0.0172)		(0.0172)
Clean winds		0.0143	0.0129
		(0.0091)	(0.0092)
Polluting winds		-0.0389 ***	-0.0372 ***
		(0.0111)	(0.0112)
Observations	17099	17099	17099

*Notes:** $p<0.1$,** $p<0.05$,*** $p<0.01$. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

B.2 Heterogeneity

B.2.1 Income

To understand the heterogeneity in the impacts of pollution on output with respect to per capita income, we group each NUTS3 region i into quantile q based on its average per capita income over 2000-2015 by defining a variable D_i^q that is equal to one if region i falls into quantile q. We then interact this group variable with pollution to recover the heterogeneous impact of pollution on output according to income quantile:

$$\Delta \log(Y_{it}) = \sum_q \beta^q D_i^q \widehat{\Delta P}_{it} + \beta_2 \Delta f(W_{it}) + \Delta\gamma_t + \nu_{it}. \tag{7}$$

We use the same instrumental variables and weather controls as in the main text, and results are reported in Figure 16.

B.2.2 Urban vs. rural

To understand the heterogeneity in the impacts of pollution on output with respect to the type of region, we group each NUTS3 region i into three groups–urban, intermediate, or rural. The classification is based on population density and proximity to urban centres and is produced by the OECD.[27] We then separately run our two-stage regression on each of these subgroups. We use the same instrumental variables and weather controls as in the main text, and results are reported in Table 11.

B.2.3 Economic structure

To understand the heterogeneity in the impacts of pollution on output with respect to economic structure, we group each NUTS3 region i as service-intensive or industry-intensive based on its

[27] See `https://www.oecd.org/regional/regional-statistics/OECD_regional_typology_Nov2012.pdf`.

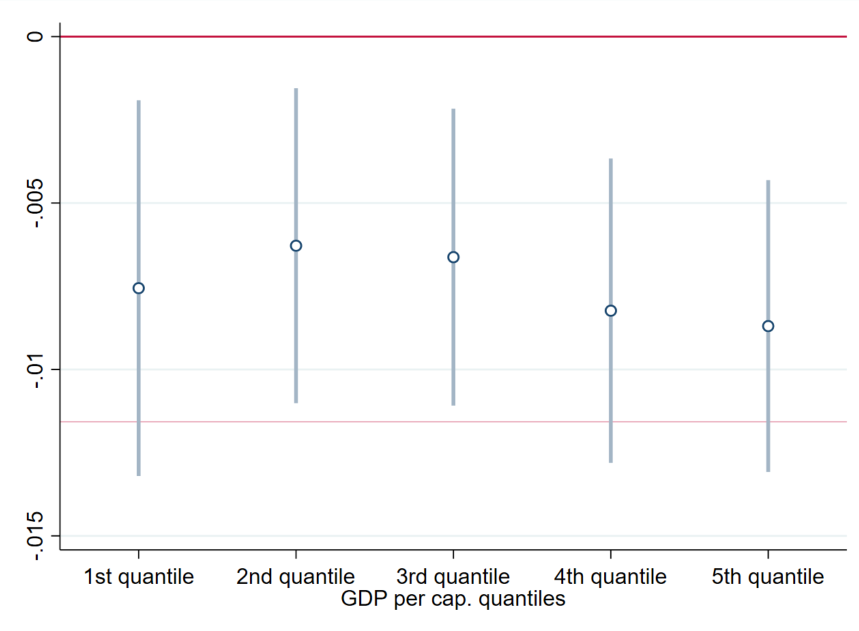

Figure 16: Heterogeneity in economic impact of pollution by income quintiles (GDP per capita). Coefficients and 95% confidence intervals from a instrumental variables regression of output per capita interacted with income quantiles on instrumented pollution. Clustered standard errors (on NUTS3 level). All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

average sector output over 2000-2015.[28] We then separately run our two-stage regression in each of these subgroups. We use the same instrumental variables and weather controls as in the main text, and results are reported in Table 12.

[28]To classify regions as industry- or service-intensive, we look at the gross value added of the industry sector divided by total GDP, and the sum of gross value added of the Professional, Information/Communication, Real Estate, and Finance sector divided by total GDP. We split each at the median. If it is above median in industry share, but below in services we classify as industry intensive. If it is above median in service share, but below in industry share we classify as service intensive. Note that this approach excludes certain regions that cannot be cleanly classified.

Table 11: Economic of pollution by region type.

	(1) Urban	(2) Intermediate	(3) Rural
$PM_{2.5}$	-0.009 **	-0.015 ***	0.003
	(0.004)	(0.003)	(0.004)
Observations	4723	7055	5321
R^2	0.088	-0.040	0.079
Weak id. stat.	36.07	43.34	24.78
Hansen J stat. p-value	0.0240	0.0108	0.0477

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

Table 12: Economic of pollution by economic structure type.

	(1) Industry intensive	(2) Service intensive
$PM_{2.5}$	-0.008 *	-0.014 ***
	(0.005)	(0.003)
Observations	2994	6518
R^2	0.069	0.044
Weak id. stat.	23.41	47.80
Hansen J stat. p-value	0.0871	0.635

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

B.3 Robustness checks

We perform a series of robustness checks to determine how our main results are affected by alternative assumptions. In all regression tables reported in this section, we report the results of the second stage of a two-stage regression. Unless otherwise reported, instruments are identical to the main text.

B.3.1 Weather controls

As emphasized above, our instrumental variables satisfy the exclusion restriction conditionally. That is, conditional on ground-level weather, both thermal inversions and wind direction should only affect economic outcomes via their effect on pollution. Because both of these variables are likely correlated with weather, which can itself impact economic outcomes, it is important to carefully control for weather. We do this in the main results using a flexible approach to controlling for temperature, precipitation, and wind speed, and including other weather variables using second-degree polynomials, and with interactions between temperature and humidity variables. In Table 13, we show that our results are invariant to adopting an even more flexible approach to including the effect of temperature (conditioning on 70 temperature bins, rather than 20 as in the baseline). Moreover, we estimate the model without conditioning on any weather covariates and continue to obtain very similar results as in the baseline specification, suggesting that our choice of weather controls is unlikely to substantially affect the results.

Table 13: Robustness with respect to weather controls.

	(1) 20 temp. +hum. int.	(2) 20 temp. bins	(3) 70 temp. bins	(4) 70 temp. bins +hum. int.	(5) no weather controls
$PM_{2.5}$	-0.0116 ***	-0.0134 ***	-0.0132 ***	-0.0106 ***	-0.0082 ***
	(0.0021)	(0.0021)	(0.0021)	(0.0022)	(0.0014)
Observations	17099	17099	17099	17099	18437
Weak id. stat.	107.0	120.5	131.6	108.0	194.9
Hansen J stat. p-value	0.242	0.0157	0.00169	0.130	0.000460

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects. Column (1) includes 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Column (2) is the same except excluding the interaction between temperature and humidity. Column (3) is the same as (2) except it includes 70 temperature bins, rather than 20. Column (4) is the same as (1) except with 70 temperature bins instead of 20. Column (5) excludes all weather controls. Regression coefficients weighted by each region's population.

B.3.2 Alternative instruments

In our main specification, we adopt (by necessity) particular definitions for the wind direction and thermal inversion variables. Specifically, we use the share of days of thermal inversions at the lowest atmospheric level in both winter and summer, and the share of days with winds from the dirtiest and cleanest three wind directions (based on pre-sample data) as instruments, for a total of four instruments. We re-run the analysis with different definitions for both of these sets of instrumental variables. Results which vary the set of inversion instruments are shown in Table 14, and results which vary the set of wind instruments are shown in Table 15.

Table 14 varies the variables used to define inversions, both based on whether and how we distinguish the season at which atmospheric inversions take place, as well as the level at which atmospheric level inversions are measured. In total, we define inversion seasonality in three alternative ways: by distinguishing winter vs. summer (as in the baseline, where we have two inversion instrumental variables corresponding to these seasons), by distinguishing four seasons (with four instrumental variables for inversions), or by not distinguishing between seasons. We also define the level of measurement for inversions in three alternative ways, as described in Section 4: at low levels (as in the baseline), at any level, or relative to surface. In total, this gives 9 possible combinations of inversion instrumental variable combinations, and results using all combinations except the baseline are provided in Table 14. Results are always negative, statistically significant, and close to the baseline results, suggesting that the particular definition of inversion chosen in not material to the results.

Table 15 varies the variables used to define wind directions. The baseline results used two wind instrumental variables, based on the number of days the wind originates from the three dirtiest and three cleanest compass octants, determined based on pre-sample data. Table 15 shows that results are not substantially affected by alternative choices of wind instrumental variables. Using only the number of days from only the cleanest directions or only the dirtiest directions, with only one wind IV in each case, delivers qualitatively similar results. Using the two cleanest and two dirtiest octants (rather than the three in the baseline) also delivers similar results, as does using seven IVs, with each capturing the share of days from one of the seven compass octants (and the eighth dropped to avoid collinearity). The final column of 15 uses an alternative approach. In this column, clean and dirty wind directions are not determined based on pre-sample data, but instead based on the prior five years of daily wind and pollution measures, such that the ranking of wind directions for each NUTS3 region is updated continually over time. This approach continues to deliver a negative and statistically significant coefficient, but somewhat smaller than the baseline results.

Overall, our results are robust to using a large set of different instrumental variables. The main coefficient varies between -0.007 and -0.013 depending on the instrument, relative to a baseline estimate of -0.011. In all cases, we find a substantial and precisely estimated impact of pollution on economic output.

Table 14: Robustness with respect to inversions instrument choice.

	(1) Inversions any (annual) b/_star/se	(2) Inversions any (2 seasons) b/_star/se	(3) Inversions any (4 seasons) b/_star/se	(4) Inversions low (annual) b/_star/se	(5) Inversions low (4 seasons) b/_star/se	(6) Inversions surface (annual) b/_star/se	(7) Inversions surface (2 seasons) b/_star/se	(8) Inversions surface (4 seasons) b/_star/se
$PM_{2.5}$	-0.0096 ***	-0.0082 ***	-0.0095 ***	-0.0117 ***	-0.0110 ***	-0.0111 ***	-0.0113 ***	-0.0116 ***
	(0.0019)	(0.0019)	(0.0018)	(0.0022)	(0.0020)	(0.0021)	(0.0021)	(0.0020)
Observations	17099	17099	17099	17099	17099	17099	17099	17099
Weak id. stat.	147.0	145.8	86.90	111.9	82.55	115.1	99.99	77.05
Hansen J stat. p-value	0.0275	0.0310	0.0976	0.123	0.245	0.102	0.191	0.308

Notes:* $p<0.1$,** $p<0.05$,*** $p<0.01$. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

Table 15: Robustness with respect to wind instrument choice.

	(1) Bad winds (Dirtiest 3 only) b/_star/se	(2) Bad winds (Cleanest 3 only) b/_star/se	(3) Bad winds (Dirtiest 2 and cleanest 2) b/_star/se	(4) All winds (7 directions) b/_star/se	(5) Rolling winds (7 directions) b/_star/se
$PM_{2.5}$	-0.0133 ***	-0.0105 ***	-0.0098 ***	-0.0101 ***	-0.0066 ***
	(0.0028)	(0.0022)	(0.0021)	(0.0020)	(0.0021)
Observations	17099	17099	17099	17099	17099
Weak id. stat.	78.47	133.5	97.73	54.20	30.06
Hansen J stat. p-value	0.619	0.949	0.490	0.0199	0.000373

Notes:* $p<0.1$,** $p<0.05$,*** $p<0.01$. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

B.3.3 Alternative specifications

We test the sensitivity of our results to the inclusion of alternative fixed effects and control variables. Our main research design controls for persistent heterogeneity between regions using a differencing approach, and controls for common shocks across Europe using year fixed effects. Table 16 replicates the main results, but adding additional controls to account for heterogeneity over time and between regions that is potentially correlated with economic activity and the instruments. Column (1) adds linear NUTS3 time trends to capture potential common trends across regions between economic activity and the instruments. Column (2) adds linear country time trends. In both cases, the coefficient remains at a similar magnitude and retains statistical significance. Column (3) adds quadratic country trends and again has little impact on results. Column (5) adds country-year fixed effects. These control for any unobserved heterogeneity at a country-year level, and identify the impact of air pollution only from within country-year variation across NUTS3 regions in (instrumented) pollution. Unfortunately, these fixed effects remove too much variation in pollution, such that standard errors increase, and the effect of pollution on output can no longer be identified precisely. As an alternative, column (4) uses country-2year fixed effects (results are identical no matter whether we start the 2-year groupings on an odd or even year). From an identification perspective, country-2year fixed effects allow nearly as much flexibility as country-year effects, but they do not soak up quite as much variation in air pollution, and the impact of air pollution on output again is precisely estimated, although only about half the magnitude as the baseline results.

In Table 17 we test the impact of controlling for additional pollutants. We focus on SO_2, because this is estimated in the MERRA-2 dataset. We now have two endogenous pollutants, which are highly correlated, in the second stage regression. To improve the predictive power of the first stage regression, we choose sets of instruments with more individual instruments from the sets of instruments we use in the main text. For example, we use four inversion instruments, which capture the share of inversions across four different seasons, and seven wind instruments, which capture wind from each of the eight compass octants. Table 17 shows the effect of both $PM_{2.5}$ and

Table 16: Regression models with alternative fixed effects.

	(1) Linear NUTS3-trends	(2) Country trends	(3) Country squared trends	(4) Country 2-year FE	(5) Country year FE
$PM_{2.5}$	-0.0110 ***	-0.0112 ***	-0.0118 ***	-0.0046 ***	-0.0041
	(0.0021)	(0.0021)	(0.0022)	(0.0016)	(0.0042)
Observations	17099	17099	17099	17097	17085
Weak id. stat.	104.2	106.0	100.0	125.7	22.56
Hansen J stat. p-value	0.423	0.438	0.695	0.0000307	0.0457

Notes:* p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

SO_2 on GDP using the two stage least squares approach, using different combinations of these instruments. Despite choosing combinations of instruments with more individual instrumental variables, our instruments are not strong at predicting both pollutants separately, and we find no effect of SO_2 on output. We continue to find a large impact of $PM_{2.5}$ on output. Although the coefficient in some specifications is larger than our baseline result, because of the large standard error it is not statistically distinguishable from our baseline estimate. We choose column (3) as our preferred co-pollutant regression as it has the largest F-statistic in the first stage (although still suggestive that our instruments are not strong enough to disentangle the effects of two pollutants).

Table 17: Regression models with controls for co-pollutants.

	(1) Instrument set 1	(2) Instrument set 2	(3) Instrument set 3
$PM_{2.5}$	-0.0181 **	-0.0124 **	-0.0119 ***
	(0.0078)	(0.0051)	(0.0045)
SO_2	0.0150	0.0046	0.0040
	(0.0167)	(0.0089)	(0.0084)
Observations	17099	17099	17099
Weak id. stat.	2.735	4.805	5.302
Hansen J stat. p-value	0.431	0.0177	0.0273

Notes:* $p<0.1$,** $p<0.05$,*** $p<0.01$. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population. Instrument set 1 includes the share of days with surface level inversions in each of four seasons as well as the share of days with winds from the three cleanest and three dirtiest compass octants. Instrument set 2 includes the share of days with low level inversions in winter and summer as well as the share of days with winds from all seven compass octants (the eighth is dropped to avoid collinearity). Instrument set 3 includes the share of days with surface level inversions in all four seasons as well as the share of days with winds from all seven compass octants (the eighth is dropped to avoid collinearity).

Table 18: Placebo estimation with randomly assigned GDP.

	(1) Placebo ln(GDP)
$PM_{2.5}$	0.002
	(0.003)
Observations	14498
R^2	0.011
Weak id. stat.	55.01
Hansen J stat. p-value	0.455

*Notes:** p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

B.3.4 Placebo test

We estimate a placebo model, or test of design, to ensure that our econometric specification is not delivering the results we are obtaining. To implement our test of design, we randomly assign pollution and meteorological variables across regions in our data set, and re-estimate the model with these placebo treatment variables. Table 18 shows that the placebo regression delivers a precisely estimated coefficient of zero on the instrumented pollution, suggesting that our results are not driven by our empirical approach, but instead by the data.

Table 19: Robustness with respect to database choice.

	(1) Van Donkelaar	(2) CAMS	(3) Monitoring stations
$PM_{2.5}$	-0.008 ***	-0.007 ***	-0.008
	(0.002)	(0.002)	(0.008)
Observations	17099	13482	2152
R^2	-0.059	0.015	-0.310
Weak id. stat.	16.76	45.30	0.975
Hansen J stat. p-value	0.00142	0.0000583	0.458

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

B.3.5 Alternative air pollution data

We re-estimate the model with different air pollution data (see section 4), based both on the Van Donkelaar et al. (2016) and the CAMS models. Results are reported in Table 19. In both cases, we continue to find a negative and statistically significant impact of pollution on GDP, although the estimates are somewhat smaller than in our baseline (but not statistically significantly so). Note that the data for CAMS includes fewer observations than for MERRA2 or Van Donkelaar et al. (2016), because of incomplete temporal coverage as described in section 4. We also include in column (3) a regression that uses data from the pollution monitoring network. Note that this network has very sparse coverage, particularly in the early years in our data set. We trim 1% of outlying observations because the distribution of monitoring station data contains some extreme values. We report an imprecise finding using this data, but the point estimate is close to that with other data sets.

B.3.6 Weighting

Our baseline estimates are population-weighted, as described in the text. We re-estimate the model without weighting, and weighting by GDP rather than population (Table 20). Unweighted estimates are slightly larger (in absolute value) than population-weighted estimates, which are slightly larger than GDP-weighted estimates. Neither of them are statistically different from the baseline, but they do suggest that the impact of air pollution on economic activity is slightly larger in small regions (as measured by GDP or population) than in larger ones.

Table 20: Economic effect of PM2.5 (weighted).

	(1) Not weighted	(2) GDP weighted
$PM_{2.5}$	-0.0129 ***	-0.0101 ***
	(0.0019)	(0.0024)
Observations	17099	17099
Weak id. stat.	141.7	93.70
Hansen J stat. p-value	0.00000217	0.222

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared.

Table 21: Economic effect of PM2.5 (without extreme values).

	(1) Outliers dropped at 1%	(2) Outliers dropped at 5%
$PM_{2.5}$	-0.0123 ***	-0.0174 ***
	(0.0021)	(0.0028)
Observations	16745	15396
R^2	0.019	0.002
Weak id. stat.	153.3	105.1
Hansen J stat. p-value	0.125	0.0284

Notes: * p<0.1, ** p<0.05, *** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

B.3.7 Dropping outliers

To ensure that our results are not driven by extreme values, we re-run the regression, but this time dropping extreme values. Specifically, we remove the observations below the 1st (or 5th) and above the 99th (or 95th) percentile of the $\Delta PM_{2.5}$ distribution. As shown in Table 21, we continue to obtain similar results and the coefficient increases somewhat in absolute value. This shows that, if anything, including outliers leads us to underestimate the impact of air pollution on economic activity.

B.3.8 Spatial autocorrelation

We test the sensitivity of our results to spatial autocorrelation and to the inclusion of spatial controls. Air pollution travels across regional borders and there are external effects (spillovers) of GDP between regions (e.g. through cross-region trade). Thus, our baseline estimation might be biased if we dont control for the potential endogeneity arising from spatial properties of pollution and economic growth.

In Table 22, columns (1) and (2) include spatial lags of pollution and GDP, where we include inverse distance weighed average pollution and GDP in a 100km and 200km radius around each regions centroid as an additional control variable. Results are not impacted by the addition of these spatial controls.

For this estimation we construct five new variables: the spatial lags of GDP, $PM_{2.5}$, and the three instruments. We calculate the distance between the NUTS3 regions centroids (within radius k) and use inverse distance weighing to construct an average value of the neighbours of the region. Thus a spatial lag of variable x for region i is defined as:

$$x_{i,SL} = \frac{\sum \frac{x_j}{d(x_i,x_j)}}{\sum 1/d(x_i,x_j)} 1(d(x_i,x_j) < k)$$

We choose two cut-off points (k values) for the indicator function ($1[\cdot]$): 100 km and 200 km. Neighbours outside of these radii are not counted in the spatial lag. This reduces the potential influence of outliers.

Finally, columns (3) and (4) implement clustering on NUTS2 and NUTS1 regions, respectively. These allow arbitrary correlation within these broader regions, over time. The coefficient remains statistically significant at the 1% level as in the baseline regression.

B.3.9 Log-log results

Our main results are presented in log-linear form. Here we present our main results in log-log form. Table 23 shows first stage results. Summer inversions cause a $\exp 0.465 - 1 = 59\%$ increase

Table 22: Robustness with spatial controls.

	(1) 100 km spatial lag	(2) 200 km spatial lag
$PM_{2.5}$	-0.005 **	-0.009 ***
	(0.002)	(0.002)
Spatial lag of $PM_{2.5}$	-0.001 **	0.000
	(0.001)	(0.000)
Spatial lag of ln(GDP)	0.003 ***	0.000
	(0.001)	(0.001)
Observations	15271	15791
R^2	0.048	0.028
Weak id. stat.	47.87	59.61
Hansen J stat. p-value	0.0472	0.0259

Notes:* p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

in pollution concentration. This is comparable to the log-linear version in Table 2, which suggests summer inversions cause a 5.9μg/m^3 increase in pollution concentrations, relative to a mean of 11μg/m^3, or 53%. Other coefficients are likewise similar in magnitude and significance.

Table 24 shows our main results in log-log form. Again, the magnitude and significance of our results are largely preserved relative to the log-linear version in the text. The log-log results indicate that a 1% increase in pollution causes a 0.12% reduction in GDP. Table 3 in the text shows that a 1μg/m^3 increase in pollution causes a 1.2% reduction in GDP. A 1μg/m^3 increase reflects a 9.1% increase from mean pollution levels in our sample, suggesting an elasticity of GDP to pollution of -0.13, very similar to the log-log results.

Table 23: First stage results in log-log form

	(1)	(2)	(3)
	$\ln(PM_{2.5})$	$\ln(PM_{2.5})$	$\ln(PM_{2.5})$
Summer inversions (low)	0.584 ***		0.465 ***
	(0.086)		(0.085)
Winter inversions (low)	0.034		0.084 *
	(0.043)		(0.044)
Clean winds		-0.252 ***	-0.243 ***
		(0.023)	(0.024)
Polluting winds		0.169 ***	0.157 ***
		(0.025)	(0.023)
Observations	17099	17099	17099
Adjusted R^2	0.224	0.262	0.274

Notes:* p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

Table 24: Main results in log-log form

	(1)	(2)	(3)
	ln(GDP)	ln(GDP)	ln(GDP)
$\ln(PM_{2.5})$	-0.1252 **	-0.1170 ***	-0.1175 ***
	(0.0515)	(0.0212)	(0.0199)
Observations	17099	17099	17099
Weak id. stat.	29.05	244.8	146.9
Hansen J stat. p-value	0.549	0.0514	0.248

Notes:* p<0.1,** p<0.05,*** p<0.01. Clustered standard errors (on NUTS3 level) are in parentheses. All estimations are conducted in first differences and include year fixed effects, 12 wind speed bins, 20 temperature bins, 20 precipitation bins, second order polynomials of relative humidity and atmospheric pressure, and interactions between temperature bins and humidity and humidity squared. Regression coefficients weighted by each region's population.

C Implications for cost-benefit analyses of air pollution control policies

Table 25: Annual market and non-market benefits from PM2.5 emission reduction scenarios in Europe.

2025 scenario (EU28)	6A	6B	6C	6D
Reduction in emissions wrt baseline	17%	25%	34%	45%
Lost working days (M€)	726	1421	2137	2831
Damage to built environment	53	106	145	162
Crop value losses	61	101	278	630
Healthcare costs	219	437	657	886
Total market benefits	1,059	2,065	3,237	4,509
Total benefits (low valuation, M€)	14,997	29,767	44,686	59,642
Total benefits (high valuation, M€)	50,317	100,937	150,853	200,074

Notes: Scenario 6A: 25% gap closure between baseline and Maximum Technically Feasible Reduction (MTFR); Scenario 6B: 50% gap closure between baseline and MTFR; Scenario 6C: 75% gap closure between baseline and MTFR; Scenario 6D: 100% gap closure between baseline and MTFR. Total benefits includes both market and non-market benefits (e.g., mortality).
Source: European Commission (2013).

Table 26: Benefits from the US Clean Air Act Amendments

Endpoint	Valuation (million 2006 US$)
Minor restricted activity days	6,700
Work loss days	2,700
Outdoor worker productivity	170
Agricultural and forest productivity	11,000
Mortality	1,800,000

Source: US EPA (2011).

Table 27: Compliance costs for $PM_{2.5}$ concentration reduction scenarios in Europe. 2008 Air Quality Directive 2008/50/EC.

2025 scenario (EU25)	Scenario A	Scenario B
Reduction in average urban background concentration of $PM_{2.5}$	-20%	-25%
Marginal abatement cost (M€/year)	4974.4	8079.6
Marginal abatement cost (€/person/year)	10	16
GDP	-0.03%	-0.06%

Source: European Commission (2008)

Table 28: Annual compliance costs associated with the US Clean Air Act Amendments

Category	Valuation (million 2006 US$)
Electric utilities	13,000
On-road vehicles and fuel	27,200
Local controls	13,500
Others	14,800
Total costs	68,500

Source: US EPA (2011)