Journal of Environmental Economics and Management 130 (2025) 103102

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

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

Air pollution and innovation☆

Felix Bracht [a], Dennis Verhoeven [b,c,d],*

[a] UBS Bank, Zurich, Switzerland
[b] SKEMA Business School, Université Côte d'Azur (GREDEG), Euralille, France
[c] London School of Economics and Political Science (LSE), Centre for Economic Performance, London, UK
[d] KU Leuven, Department of Managerial Economics, Strategy and Innovation (MSI), Leuven, Belgium

ARTICLE INFO

JEL classification:
O30
Q53
O13

Keywords:
Air pollution
Air quality
Innovation
Productivity

ABSTRACT

If air pollution harms innovation — and therefore future productivity — existing assessments of its economic cost are incomplete. We estimate the effect of fine particulate matter concentration on inventive output in 977 European regions. Exploiting thermal inversions and weather-induced ventilation of pollutants for identification, we find that a decrease in air pollution equivalent to the average yearly drop in Europe leads to 1.2% more patented inventions in a given region. A back-of-the-envelope calculation suggests that accounting for the effect on innovation increases the economic cost of air pollution as assessed in prior work by about three quarters.

1. Introduction

Quantifying the economic cost of air pollution is important for rational abatement policies. Obtaining a complete picture of these costs is challenging as poor air quality affects productivity through multiple channels (Zivin and Neidell, 2018; Aguilar-Gomez et al., 2022). Hospitalization and mortality statistics can be used to calculate the forgone output due to the most severe health impacts. For instance, Landrigan et al. (2018) find that the cost of air pollution due to premature deaths amounts to 0.092% of GDP globally, with developing countries most affected. Such assessments do not include the harm of less acute health and cognitive effects. These so-called 'sub-clinical' effects have been found to reduce the supply and productivity of labor, even in occupations demanding mainly mental effort (Zivin and Neidell, 2012; Hanna and Oliva, 2015; Chang et al., 2016, 2019). Aggregating across occupations, countries, and productivity mechanisms, Dechezleprêtre et al. (2019) find that a 10% improvement in air quality increases GDP in the same year by 0.8%. Focusing on US counties, Avila Uribe (2023) find smaller but economically important effects for rural areas and specific economic sectors.

Estimating the aggregate effect of air pollution on same-year GDP is attractive because it nets out a range of underlying mechanisms through which economic output can be affected. Doing so establishes the external validity of productivity effects found in specific contexts. The approach, however, does not capture the economic costs due to decreased innovation because the benefits of an innovation today typically accrue over several years or even decades and will not show up in same-year GDP (Scherer, 1965;

☆ We thank Dirk Czarnitzki, Hans Degryse, Reyer Gerlagh, Stephen Machin, Ralf Martin, Tauhidur Rahman, Thomas Schaper, Gregor Singer, Catherine Thomas, Steven Vanhaverbeke, Anna Valero, John Van Reenen and Jesse Wursten for valuable comments and feedback. The resources and services used in this work were provided by the VSC Flemish Supercomputer Center (https://www.vscentrum.be/), funded by the Research Foundation - Flanders (FWO) and the Flemish Government. Verhoeven gratefully acknowledges funding from the Research Foundation Flanders (FWO grant 1241520N) and the H2020 Framework Programme (Marie Skłodowska-Curie Action DIV_INV grant 844254).

* Corresponding author at: SKEMA Business School, Université Côte d'Azur (GREDEG), Euralille, France.
E-mail addresses: felix.bracht@outlook.com (F. Bracht), dennis.verhoeven@skema.edu (D. Verhoeven).

https://doi.org/10.1016/j.jeem.2024.103102
Received 27 June 2023
Available online 1 January 2025

Ravenscraft and Scherer, 1982). Because innovation is the key driver of growth (Aghion and Howitt, 1990; Romer, 1990; Jones and Summers, 2020), this means that the economic costs of air pollution could be considerably understated when overlooking its effect on innovation.

This paper provides evidence of the causal effect of air pollution on inventive output in Europe. Our identification strategy borrows from existing literature by relying on two weather phenomena that provide plausibly exogenous variation in air pollution (Hanna and Oliva, 2015; He et al., 2019; Dechezleprêtre et al., 2019; Fu et al., 2021; Chen et al., 2022; Avila Uribe, 2023). The first instrument measures a region's exposure to 'thermal inversions' in a given year. Inversions 'trap' pollutants near the earth's surface by reversing the negative temperature-height relationship. The second instrument measures the level of natural ventilation that disperses pollutants. It is a function of wind speed and the 'planetary boundary layer' height. This layer determines the volume over which turbulent air flows diffuse pollutants. Both instruments are determined by idiosyncratic atmospheric conditions plausibly exogenous to innovative activity. One concern is that the weather in general is determined by the same conditions and also affects innovation, resulting in a violation of the exclusion restriction. To mitigate this concern, we include controls for wind speed, precipitation, temperature, humidity and air pressure. We implement our empirical strategy by collapsing data on pollution, weather conditions, and patents into a yearly panel of in 977 European NUTS-3 regions (the European equivalent to US counties) between 2001 and 2012. We measure inventive output as the total number of patents filed in the year following a given level of exposure to air pollution. Our measure of air pollution is the exposure to particulate matter with a diameter of 2.5 micrometers or less ($PM_{2.5}$).[1] It is considered one of the primary pollutants by the World Health Organization and is widely used to study the effects of air pollution.

The results from fixed effects Poisson regressions with a control function that incorporates our instruments show that a 0.17 μg/m^2 decrease in $PM_{2.5}$ concentration – the average yearly reduction in Europe over the period studied – leads to 1.2% more patented inventions, which corresponds to 1.51 patents for the average region. Total patenting in Europe between 2001 and 2012 would have been 1.94 percent lower if air pollution had remained at 2001 levels. The instruments correlate well to fine particulate matter, with F-statistics above 40, and induce considerable within-region variation for most of Europe. Weighting for patent quality increases the size of the coefficient. Our effect is driven almost exclusively by regions with above-median levels of air pollution. Further, we find a negative and statistically significant effect of air pollution in regions with high and intermediate levels of urbanization, while the effect for rural areas is also negative but not precisely estimated. Abatement benefits appear largest in urban areas, which are relatively polluted and innovation-intensive.

Our baseline results measure the effect of air pollution on innovation output for the average region. One concern with aggregating these region-level estimates to economy-wide costs is that they may result from inventors who relocate rather than decreased total output. If inventors migrate in response to air pollution, innovation lost in one region may be partially recovered in another one. We rule out this mechanism using address information across inventors' subsequent patent applications to identify migration between regions. We find no statistically or economically significant effect of air pollution on inventor in- or out-migration.[2]

A key assertion of this paper is that the effect on innovation is not accounted for in prior assessments that estimate effects on same-year GDP. In further analyses, we address two concerns about interpreting our results as a 'hidden cost' of air pollution. First, firms may reduce their R&D spending in response to air pollution, for instance because they expect tighter regulations that shrink future profits. Because R&D spending is part of a region's GDP, the damages to innovation through this channel would be partially accounted for by the estimates of Dechezleprêtre et al. (2019). We argue that R&D budgets are long-term decisions that are unlikely to change much with year-to-year variation in weather-induced air pollution. Furthermore, our finding that low-quality projects are not more affected is inconsistent with the R&D spending mechanism which would suggest low-quality projects to be discontinued first. A second concern is that, if inventive output goes down because of increased (inventor) mortality, traditional cost estimates based on productive life years lost in the population would partly account for our main result. We find, however, that mortality is only marginally affected in our set-up, and cannot explain the effect on innovation.

The paper's final part quantifies the costs implied by lost innovation with a back-of-the-envelope calculation. The challenge here is to approximate the total value an innovation creates over its lifetime. We rely on several estimates from the literature to assess the private value (Deng, 2007; Bessen, 2009; Gambardella et al., 2008) and spillovers created (Bloom et al., 2013; Zacchia, 2020; Myers and Lanahan, 2022; Arque-Castells and Spulber, 2022) by patented inventions. While the result of this exercise is subject to considerable uncertainty because value estimates in the literature do not agree well, the cost implied by the effect on innovation is large. Under our preferred assumptions, the yearly cost of our reference increase in air pollution due to missed innovations amounts to 0.11% of GDP, about 0.78 times the cost due to a reduction in same-year economic output as estimated in Dechezleprêtre et al. (2019).

This paper relates to the burgeoning literature on the productivity effects of air pollution. The idea that environmental quality affects productivity is well-established in environmental economics. Theoretical work has used endogenous growth models to examine under what conditions sustainable economic growth is feasible when production suffers from environmental damage (Bovenberg and Smulders, 1995, 1996). A large body of evidence supports this assumption. Early assessments leverage clinical data to document important effects of air pollution on mortality and morbidity (see Landrigan et al., 2018, for an overview). More recent empirical work sheds light on more subtle, 'sub-clinical' health effects (see Aguilar-Gomez et al., 2022, for a recent review). Labor supply

[1] The fact that these particles are so small, allows them to enter the bloodstream through the lungs and affect the functioning of several organs (Underwood, 2017).

[2] Prior work has established that location choices depend on air quality (Khanna et al., 2021; Heblich et al., 2021). We rationalize our opposing finding by noting that the overall level of air pollution or the variation induced by our instruments is too small to invoke a migration response.

decreases in response to polluted air at the time of exposure in samples of households living in Mexico City and Lima (Hanna and Oliva, 2015; Aragón et al., 2017). Using the 1970 Clean Air Act as a source of variation, Isen et al. (2017) find evidence of effects on long-run labor force participation and earnings. In addition to these extensive margin effects, a series of studies have found negative effects on the intensive margin of labor productivity for physically demanding occupations such as agricultural work, pear packing, and professional soccer (Zivin and Neidell, 2012; Chang et al., 2016; Lichter et al., 2017). Further work has extended these findings to more cognitively demanding settings. Exposure to air pollution reduces performance of students taking exams, baseball umpires, call center workers, chess players, investment analysts, and individuals playing brain-training games (Stafford, 2015; Ebenstein et al., 2016; Archsmith et al., 2018; Chang et al., 2019; Künn et al., 2023; Dong et al., 2021; La Nauze and Severnini, 2021; Duque and Gilraine, 2022; Cook et al., 2023). Aggregating over the various mechanisms through which air quality affects economic output, economically relevant effects of air pollution on same-year GDP have been uncovered (Dechezleprêtre et al., 2019; Avila Uribe, 2023). Our paper shows that the costs of air pollution likely extend beyond such immediate losses to GDP because reduced innovation will harm future (productivity) growth.

The remainder of the paper proceeds as follows. Section 2 describes the data construction and presents descriptive findings. Section 3 outlines the empirical strategy we employ to estimate the causal effect of air pollution on inventive output. Section 4 discusses the results of our preferred specification, documents several robustness checks, and explores heterogeneity across regions. Section 5 provides the evidence to rule out mechanisms related to inventor mobility, R&D expenditures, and mortality. Section 6 provides the details of our cost calculation. Section 7 concludes.

2. Data

2.1. Data sources

Patents. To measure inventive output, we rely on patent data from the EPO PATSTAT database (Spring 2018 edition). We link patent filings to NUTS-3 regions in Europe using the geolocation effort described in de Rassenfosse et al. (2019). This method retrieves geographic coordinates from inventor addresses using online geolocation services. An invention is typically described in various patent documents at different times. To count innovations, we use so-called patent families, which group all patent documents related to the same invention.[3] Because one invention can be linked to multiple published documents, one inventor may be linked to multiple addresses. The geolocation method assigns geographical coordinates to the inventor's address disclosed first. When tracking inventor migration, we use data that allows tracing inventor addresses over time (Morrison et al., 2017). This method assigns a stable inventor identifier across multiple patent documents for patents filed at the European Patent Office (EPO), the World Intellectual Property Office (WIPO), and the US Patent and Trademark Office (USPTO). This dataset is appropriate for identifying inventor relocation for two reasons. First, it allows tracking individual inventors across patent applications by employing a name disambiguation algorithm. Second, it geolocates addresses from each publication associated with the patented innovation, rather than the priority filing only, as is done in de Rassenfosse et al. (2019). These features allow us to detect inventor relocation events more precisely because there are more documents from which we can infer an inventor's address. When assessing patent quality, we use the number of times a patent was cited. This indicator is extracted from the PATSTAT database and counts the number of patented inventions that cite the focal invention.[4]

Our sample consists of 1,415,353 patent families with a first filing between 2001 and 2012. The data from de Rassenfosse et al. (2019) provide a geocode for inventions with at least 1 patent document published by 2016. We cut the sample in 2012 to avoid truncation bias due to a drop-off in patent counts during recent years of the database. As a patent application only enters our data set when the first patent document is published, the time series towards the end of the database shows a negative trend. This drop in patent applications is artificial because an increasing portion of patent applications have not produced a published document that ends up in the database. To avoid truncation bias due to the fact that publication lags could correlate to unobserved determinants of innovation, we restrict our sample to the period 2001 until 2012, before truncation starts.

Pollution. We use gridded data on the annual average ground-level fine particulate matter ($PM_{2.5}$) concentrations from Hammer et al. (2020) and Van Donkelaar et al. (2019). These data contain estimated pollution concentrations at a fine spatial resolution of $0.01° \times 0.01°$ by combining Aerosol Optical Depth (AOD) retrievals from satellites with regional ground-level observations. We combine pollution data with gridded data on the population of Europe from CIESIN (Columbia University, 2018) to obtain the population-weighted $PM_{2.5}$ concentration.[5]

Meteorological information. We obtain atmospheric data from the MERRA-2 dataset (Gelaro et al., 2017) to determine the presence of thermal inversions. This data set contains daily mean temperature and relative humidity for 42 different pressure levels

[3] The patent application process typically takes several years. After filing, examiners produce an initial search report documenting the relevant prior art and assessing patent criteria such as novelty and inventive step. In response to this search report, applicants are often required to modify the patent application before it is granted. At each of these stages in the process, patent offices publish patent documents that we observe in the database. In addition, innovators seeking protection across different jurisdictions will file several patent applications at different patent offices. The PATSTAT database groups all documents related to an invention into a so-called patent family.

[4] It uses the concept of patent families mentioned before to identify the invention covered by different patent documents. This approach avoids double-counting citation linkages between patent documents representing the same invention.

[5] Population estimates are available every five years starting in 2000 and ending in 2020. We always use the closest population estimated to calculate annual weighted $PM_{2.5}$ concentration.

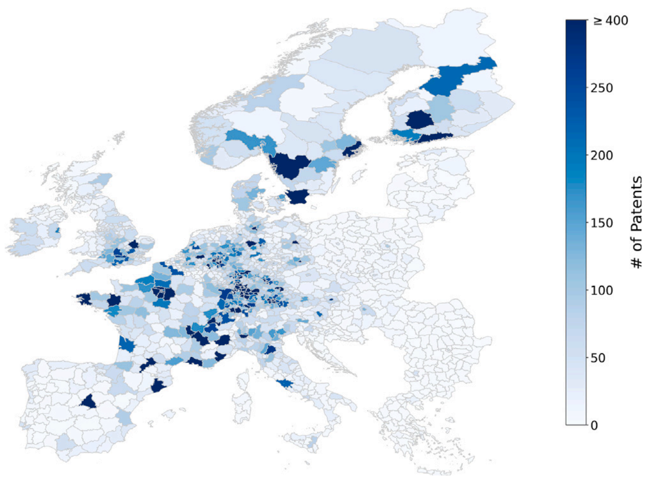

Fig. 1. Patent intensity by NUTS-3 region.
Notes: Plot of the annual average number of patent filings per NUTS-3 region. Darker blue regions have a higher patent intensity. Patent data are from the 2018 Spring edition of PATSTAT, and are assigned to NUTS-3 regions using the geocoding of inventor addresses as described in de Rassenfosse et al. (2019). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

at respective heights. This information allows us to calculate the temperature-height relationship that defines thermal inversions. Following Dechezleprêtre et al. (2019), we identify the occurrence of a thermal inversion when the temperature is higher at the second-to-lowest level of the atmosphere than at the lowest level. To circumvent the problem that the spatial resolution ($0.5° \times 0.625°$) of these data is too coarse to cover smaller NUTS-3 regions, we interpolate these data to a finer resolution ($0.01° \times 0.01°$). For temperature we use a simple bi-linear interpolation technique because the temperature varies between geographic locations rather continuously. As relative humidity is spatially discontinuous, we apply the nearest-neighbor interpolation technique. The thermal inversion instrument is calculated as each region's annual share of days with a thermal inversion.

Further atmospheric and climate data for the period 2001–2018 come from the family of ERA5 datasets provided by the European Centre for Medium-Range Weather Forecasts.[6] Specifically, we make use of the ERA5 data (Hersbach et al., 2020) and the ERA5-Land data (Muñoz-Sabater, 2019) to calculate the ventilation coefficient and additional weather control variables. We use ERA5 hourly data on the boundary layer height to calculate daily means for each NUTS-3 region.[7] We calculate daily means for temperature, wind speed, and precipitation using ERA-5-Land hourly data.[8] Ventilation Capability – our second instrument – is calculated as the annual average number of consecutive days in which the ventilation coefficient is equal or below the 20th percentile of the distribution of ventilation coefficients within a NUTS-3 region. The ventilation coefficient is calculated as the product of the daily mean wind speed and the daily mean boundary layer height.

Geographic indicators. We collect geographic indicators for NUTS-3 regions from Eurostat.[9] We use yearly data on population,[10] regional classifications[11] and mortality.[12]

2.2. Descriptive findings

Fig. 1 maps the annual average number of patent filings for the sample period by NUTS-3 region. In case multiple regions are assigned to one patent because different inventors on the patent live in different regions, we assign the patent to both regions. Western Germany, southern UK, northern France, and the Nordic countries have the highest patent output, while Eastern European countries and Spain show lower patenting activity. The spatial distribution maps well to measures of R&D intensity and R&D growth reported in the Research and Innovation Performance report published by the (European Commission, 2020), lending support to using patent filings as an indicator for inventive output.

Fig. 2 (blue line) documents a decreasing trend in $PM_{2.5}$ concentration over time. On average, $PM_{2.5}$ concentration falls by 0.17 $\mu g/m^3$ each year. A similar trend appears for nearly all air pollutants in Europe (Koolen and Rothenberg, 2019), indicating

[6] An overview of all ERA5 datasets can be found on https://confluence.ecmwf.int/display/CKB/The+family+of+ERA5+datasets.
[7] Again, we use bi-linear interpolation to increase spatial resolution from $0.25° \times 0.25°$ to $0.05° \times 0.05°$.
[8] These data are provided at a spatial resolution of $0.1° \times 0.1°$, which is sufficient to cover smaller NUTS-3 regions.
[9] Eurostat database are accessible on https://ec.europa.eu/eurostat/data/database.
[10] We use the "demo_r_pjanaggr3" dataset.
[11] We use the classification of metropolitan regions which can be found on https://ec.europa.eu/eurostat/web/rural-development/background. These data are the same as used in Dechezleprêtre et al. (2019)
[12] We use the "demo_r_deaths" dataset.

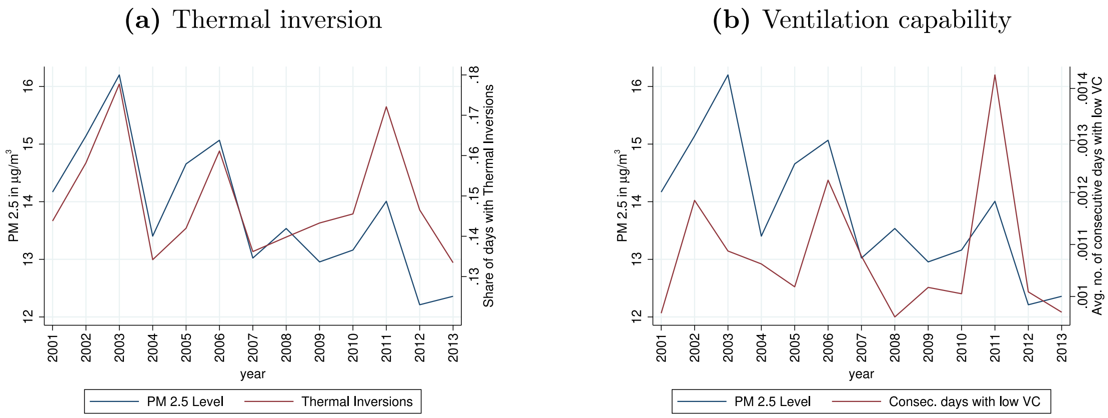

Fig. 2. Trends air pollution and instruments over time.
Notes: Compares the trend in air pollution and the instrumental variables – thermal inversions (left) and ventilation capability (right) – over time across Europe. In both graphs, the blue line shows the annual average population-weighted $PM_{2.5}$ level. The red lines show the mean value of the instrumental variables across Europe between 2001 and 2013. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

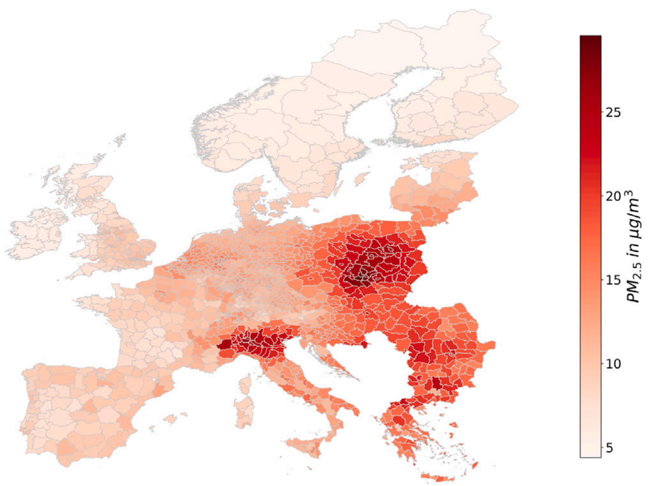

Fig. 3. Pollution levels by NUTS-3 region.
Notes: Annual average $PM_{2.5}$ level per NUTS3-region. Average annual fine particulate matter concentrations ($PM_{2.5}$) are available at a resolution of 0.01°×0.01° (Hammer et al., 2020; Van Donkelaar et al., 2019). We use population data from CIESIN (Columbia University, 2018) to obtain the population-weighted average pollution level for each NUTS-3 region.

a steady improvement in Europe's air quality. The figure also shows the correlation over time between $PM_{2.5}$ concentration and the two instrumental variables. Both thermal inversions and ventilation capability are strong predictors of $PM_{2.5}$ concentration, providing non-parametric evidence for the strength of our instruments.

Fig. 3 shows the annual average $PM_{2.5}$ concentration of NUTS-3 regions in the sample period. Eastern European regions and northern Italy experience higher pollution levels than the rest of Europe. The spatial distribution of $PM_{2.5}$ concentration maps well to pollution levels from local stations published in the Air Quality in Europe report by the European Environment Agency (2020). The variation reported here validates our measure but is not used for identification as the region-fixed effects in our specification purge all cross-sectional variation in innovation and pollution.

Our sample includes observations for 977 NUTS-3 regions in 22 countries from 2001 until 2012. We winsorize our measure of $PM_{2.5}$ concentration at 1% at each side of the distribution to minimize the influence of outliers, which imply unrealistically high or low pollution levels.[13]

Table 1 shows summary statistics for the sample data. On average, we observe 123 patent filings per year within a NUTS-3 region. This distribution is right-skewed, with a median of only 46 filed patents. The annual average population-weighted $PM_{2.5}$ concentration per NUTS-3 region amounts to 13.78.

[13] No outliers were detected for the other variables in our models.

Table 1
Summary statistics.

	Mean	Std. Dev.	10th	50th	90th
Nb. patent filings	123.11	244.05	1	46	301
Nb. fw. cites	593.01	1304.69	2	169	1468
Nb. 5 year fw. cites	199.06	425.66	0	61	487
Nb. fw. cites normalized	44.13	87.74	0.15	15.01	110.72
Nb. fam. members	585.50	1191.64	4	190	1438
Nb. triadic patents	16.76	34.41	0	5	45
Nb. EPO filings	59.80	111.04	0	22	149
Nb. EPO grants	32.29	59.85	0	12	82
Nb. filings with any grant	75.97	146.38	1	29	183
Population (*in thd*)	371.43	443.95	94.34	249.60	727.87
Pop-weighted PM 2.5	13.78	4.37	8.91	13.08	19.89
Nb. thermal inversions	54.52	33.55	11	52	101
Share thermal inversions	0.15	0.09	0.03	0.14	0.28
Ventilation Coefficient (m^2/s)	1509.99	711.21	546	1494	2393
Avg consecutive days with low VC	0.00	0.0004	0.0008	0.0010	0.0015
Relative humidity (%)	0.72	0.07	0.63	0.73	0.80
Temperature (C)	10.26	2.78	7	10	14
Wind speed (m/s)	2.33	0.87	1.16	2.29	3.40
Precipitation (m)	0.13	0.04	0.09	0.12	0.18
N	11724				

Notes: Baseline estimation sample summary statistics of 977 NUTS-regions for 22 European countries observed over the period 2001–2012.

3. Empirical strategy

3.1. Measures for air pollution and innovation

Air pollution. We proxy exposure to air pollution using the annual average $PM_{2.5}$ concentration in a region. $PM_{2.5}$ refers to suspended particulates smaller than 2.5 μm in aerodynamic diameter and is one of the priority pollutants under EU air quality regulations (European Parliament and Council of the European Union, 2008) and the National Ambient Air Quality Standards (NAAQS) in the United States (US EPA (U.S. Environmental Protection Agency), 1997). Other common air pollutants, such as nitrogen oxides and ozone, are not recorded at a scale required for our analysis. As a result, we cannot disentangle the effects of $PM_{2.5}$ from those of other correlated pollutants. We view this as only a small caveat for three reasons. First, $PM_{2.5}$ is considered the most harmful pollutant to cognitive function and physical health, limiting the confounding effect of co-pollutants. Second, $PM_{2.5}$ and other pollutants correlate because they result from similar combustion-related industrial activities. As such, regulations targeting $PM_{2.5}$ reduce co-pollutants as well, indicating that the main policy implications of this study hold even if co-pollutants partially drive the effects. Third, $PM_{2.5}$ is a key reference indicator to track air pollution globally (Power et al., 2016; WHO, 2016), and is widely used in prior work. Using $PM_{2.5}$ to track air pollution ensures comparability of our results with the largest possible number of prior studies.

Innovation. To proxy innovation output, we use the yearly number of patent families filed for in a region. Patent data provide a window on innovative output widely used in studies of technological change, innovation, and economic growth (Griliches, 1990; Schmookler, 1966; Trajtenberg, 1990). In this study, we benefit from the fact that most jurisdictions require inventors to state their addresses when applying for a patent. Recent efforts to geo-locate these addresses allow us to assign patented inventions to NUTS-3 regions.[14]

Two issues regarding patent data deserve attention here. First, not all inventions are patented because, when it is easy to maintain secrecy, the innovator has the incentive to minimize knowledge leaks to competitors. Hence, only a subset of innovations are patented. As such, a potentially substantial portion of innovation outputs are not captured by our analyses, making the estimated economic costs a lower bound estimate. Second, patents do not provide information on inventive inputs, meaning we cannot disentangle labor supply (hours worked) from labor productivity (output conditional on hours worked). Prior work has found effects of air pollution on both labor supplied and labor productivity (for instance, Hanna and Oliva, 2015 and Chang et al., 2019). Given this data limitation, our results reflect the combined effect, which we believe is of key importance when interested in quantifying the economic costs of air pollution.

3.2. Instrumental variables

Thermal inversions. Typically, temperature decreases with height in the lower part of the atmosphere (the troposphere). Less dense, warm air moves upwards and dilutes pollutants vertically. Thermal inversions reverse this monotonic relationship.

[14] NUTS-3 regions can be seen as the equivalent to US counties, and typically have a population between 125,000 and 800,000

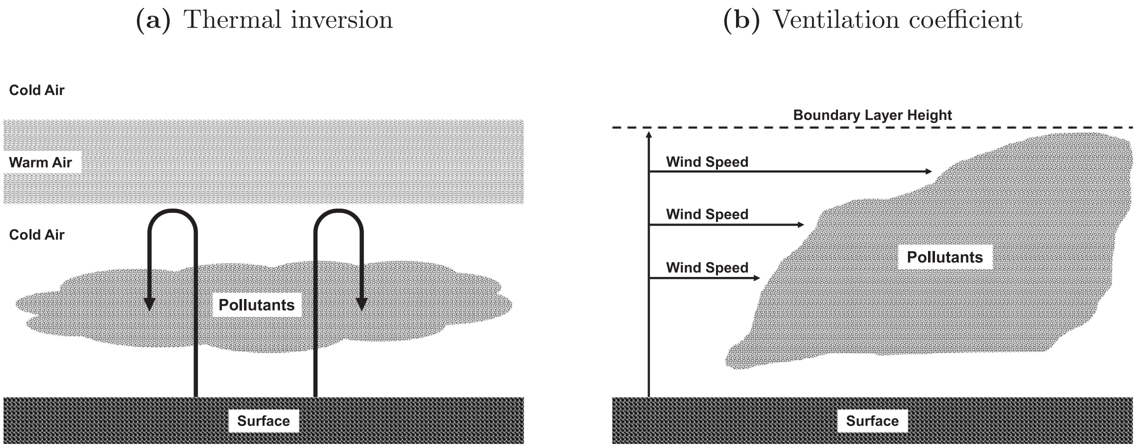

Fig. 4. Graphical representation instruments.
Notes: Graphical illustration of a thermal inversion (left) and the ventilation coefficient (right). A thermal inversion occurs when the typically negative height-temperature relation (colder at higher altitudes) is reversed. A layer of warm air obstructs air from the surface to rise, trapping pollutants close to the earths surface. The ventilation coefficient is the product of wind speed and the height of the planetary boundary layer. In this layer, turbulent air flows disperse pollutants. When the height of the layer is large, pollutants are dispersed over a larger volume of air. When there is a lot of wind, pollution is more easily dispersed.

Warmer air at higher altitudes hampers vertical dilution and traps pollutants at ground level (see Fig. 4(a)) (Vallero, 2014). Following Dechezleprêtre et al. (2019) and Arceo et al. (2016), we exploit this phenomenon to construct an instrumental variable for air pollution by calculating the annual share of days with a thermal inversion in a region.

Ventilation capability. Our second instrument measures variation in the natural ventilation of pollutants. The atmospheric literature measures the dispersion speed of pollutants using the ventilation coefficient, i.e., the product of wind speed and boundary layer height (Holzworth, 1967, 1972). Wind speed determines the horizontal dispersion of pollutants whereas boundary layer[15] height determines the volume of air among which pollutants are dispersed in the atmosphere. On days with a high ventilation coefficient, the concentration of air pollutants is low because natural ventilation allows them to dilute quickly. Fig. 4(b) provides a schematic representation of the ventilation coefficient. To construct our instrument, we first calculate the daily ventilation coefficient for each region. We then obtain the annual average length of spells of consecutive days with a low ventilation coefficient — which we call *Ventilation Capability*. The logic behind this approach is that spells of consecutive days with limited natural ventilation – rather than the average number of days – drive pollution concentrations. To clarify this, consider two regions identical in pollution generation but facing a different daily ventilation coefficient distribution. In one region, low ventilation coefficient days are followed by high ventilation coefficient days. In the other region, low ventilation coefficient days are concentrated in time. Any one day's pollution in the former region will only affect humans living there for one day before ventilation disperses it. In comparison, in the latter region, one day's pollution will affect humans for multiple days in a row before being dispersed. As a result, the average air pollution in the latter region will be higher.

Relevance assumption. Several atmospheric studies have documented a positive correlation between thermal inversions and air pollution (Gramsch et al., 2014; Xu et al., 2017). Similarly, the negative correlation between the ventilation coefficient and air pollution is well-established (Hou et al., 2018; Genc et al., 2010). Our descriptive findings (Section 2.2) and first-stage results (Section 4) show a strong correlation between both instrumental variables and air pollution, justifying the relevance assumption of our IV approach.

Exclusion restriction. Our instruments plausibly affect inventive output only through their effect on air pollution. They provide a source of exogenous variation because they result from large-scale atmospheric dynamics unrelated to human activity (Stull, 1988; Garratt, 1994). As argued in Dechezleprêtre et al. (2019), the relevant literature provides no evidence of pollution causing thermal inversions rather than vice versa. Nowhere in the meteorological literature can we find a claim that the ventilation coefficient is human-induced. One potential concern might be that our instruments correlate to other weather conditions that affect innovation through, for instance, traffic conditions or the mood of inventors. We control for average wind speed, precipitation, temperature, and relative humidity to mitigate this concern. Being exogenous themselves, they account for a potential correlation between innovation and weather conditions, which would violate the exclusion restriction. Our results show that these weather conditions have, at most, a marginal effect on innovation output in our specification.

[15] The planetary boundary layer is the lower part of the troposphere that experiences the most turbulent air flow (which is crucial for diluting pollution) because of its interaction with the surface of the earth. The height of this layer varies based on the earth's topography and meteorological conditions.

3.3. Econometric model

We employ a count data model (rather than a linear model) because the number of patent applications only has non-negative values and its distribution has a relatively large mass on low values. As a result, count models could improve the specification of the error term distribution. Count data models with multiple fixed effects that also correct for endogenous regressors have been recently established (Lin and Wooldridge, 2019). Specifically, we use a Poisson fixed-effects model, which is robust to under- and overdispersion (Wooldridge, 1999). Hence, we assume that the annual number of patent applications per NUTS-3 region follows a Poisson distribution. We use a control function approach to account for the endogeneity of the $PM_{2.5}$ concentration in our non-linear fixed-effects model. This is important because the regular two-stage least squares estimator might not be consistent in combination with a non-linear second-stage regression. The control function approach predicts residuals from the first-stage regression and allows to consistently estimate the effect of (endogenous) $PM_{2.5}$ levels in the Poisson-fixed-effects regression. The idea here is that predicted residuals from the first-stage regression will control for endogeneity.[16] The first-stage regression – i.e. the control function (CF) – used to predict residuals is given by:

$$P_{it} = \alpha_1 TI_{it} + \alpha_2 VC_{it} + \alpha_3 Pop_{it} + \alpha_4 X_{it} + \omega_{ct} + \eta_i + \sigma_i s_{it} + v_{it} \tag{1}$$

P_{it} is the average population-weighted $PM_{2.5}$ concentration for region i in year t. Following Dechezleprêtre et al. (2019), we weight $PM_{2.5}$ concentration for population density to measure individuals' exposure to air pollution rather than air pollution itself. The first instrumental variable TI_{it} measures the annual share of days with thermal inversions in region i. The second instrumental variable VC_{it} measures the average length of consecutive days with a low ventilation coefficient in year t for region i. We define a low ventilation coefficient as any daily average ventilation coefficient equal or below the 20th percentile of the distribution of daily average ventilation coefficients in region i. Pop_{it} is the total population for region i in year t. X_{it} is a vector of weather covariates controlling for weather conditions on the ground in region i. It includes annual means for wind speed, temperature, relative humidity, and total precipitation. This vector also includes the second-degree-polynomials of all weather variables to account for a more flexible functional form. ω_{ct} are country-year fixed effects, η_i are region fixed effects and s_{it} are region-specific time trends. The error term v_{it} (which we allow to be clustered at the regional level) captures all unobserved determinants of pollution.

The second-stage Poisson fixed-effects regression is given by:

$$E\left(Y_{it} | P_{it}, \widehat{v_{it}}, Pop_{it}, X_{it}\right) = exp\left(\xi_1 P_{it} + \xi_2 \widehat{v_{it}} + \xi_3 Pop_{it} + \xi_4 X_{it} + \lambda_{ct} + \kappa_i + \alpha_i s_{it} + \varepsilon_{it}\right) \tag{2}$$

Y_{it} represents the output from inventive activity in year t and is measured by the number of patent filings in the following year by inventors living in region i. P_{it} is the actual $PM_{2.5}$ concentration in year t in NUTS-3 region i. $\widehat{v_{it}}$ are predicted residuals obtained from the first-stage regression and account for the endogeneity of P_{it}. λ_{ct} are country-year fixed effects that account for country-specific time-varying effects such as exposure to macroeconomic shocks or changes in national regulations. κ_i and s_{it} are respectively region fixed effects and region-specific slopes that eliminate region-specific time-invariant effects and linear time trends that may correlate to innovation and pollution.[17] As such, our specification relies on within-region year-to-year off-trend variation in pollution caused by the instruments.[18] The error term for the second stage ϵ_{it} (allowed to be clustered at the regional level) and captures all unobserved determinants of innovation output.

Our decision to use patent filings in the year after exposure to air pollution reflects an assumption we need to make because we only observe the timing of a patent filing rather than the timing of inventive activity. We assume the bulk of inventive activity occurs during the calendar year preceding the patent filing. This lag is included to account for the fact that preparing and filing for a patent takes an estimated 3–9 months (Scherer, 1965). Also, inventors are unlikely to postpone patent filing until much after the conclusion of the inventive activity. It is in the inventors' best interest to file for a patent soon after the invention date to avoid others filing a patent for the same idea (Griliches, 1990).[19] Interviews with a patent attorney, an EPO official, and a patent data consultant confirm the 3–9 months lag is a reasonable assumption. We explore different assumptions on the lag between exposure and patent filings, and find small and insignificant effects with a zero-, two-, three-, and four-year lag.

A second timing consideration is that innovation projects vary in length. One patent in our sample might result from a 1 year project, while another could stem from a 3 year project. In our design, we estimate the effect of air pollution exposure during the last year of a project before a patent is filed. For a 1 year project, this correctly measures yearly innovation output. However, for a 3 year project, we attribute the entire patent to the project's final year. Assuming the share of projects with different duration is orthogonal to our IVs, this approach is not problematic: for each 3 year project where we assign triple the yearly innovation output, there are two other 3 year projects that we miss because they are in their first or second year of development. If patents from longer-term

[16] A simple endogenous model can be written as $y_1 = \beta_1 y_2 + \beta_3 z_1 + u_1$ where y_2 is the endogenous explanatory variable and z_1 a vector of exogenous regressors. The linear projection of y_2 on the exogenous regressors is $y_2 = \alpha_1 z_1 + v_2$ and can be the first stage in a 2SLS approach. Endogeneity of y_2 arises only if u_1 is correlated with v_2 which can be formalized as $u_1 = \rho_1 v_2 + e_1$. Hence, predicting residuals from first stage $v_2 = y_2 - \alpha_1 z_2$ and plugging them in the first stage $y_1 = \beta_3 z_1 + \beta_1 y_2 + \rho_1 \hat{v}_2 + error$ controls for endogeneity.

[17] In theory, we could control for NUTS-2 year fixed effects, but this specification absorbs a great deal of variation in the instruments, making them weak.

[18] In theory, the analysis could be performed at the level of the firm, allowing for firm-fixed effects. However, obtaining high-quality data linking patents to European firms would require a significant data collection effort. We judged this effort to be disproportionate to the potential benefit, as our identification strategy primarily relies on fixed effects to improve estimate precision rather than to identify exogenous variation.

[19] Indeed, the requirements for an invention to qualify for a patent are novelty and non-obviousness. Any evidence that the invention was already known before, thus, violating the novelty criterion, is called prior art. Consequently, the closer the filing and invention date, the smaller the pool of potential prior art.

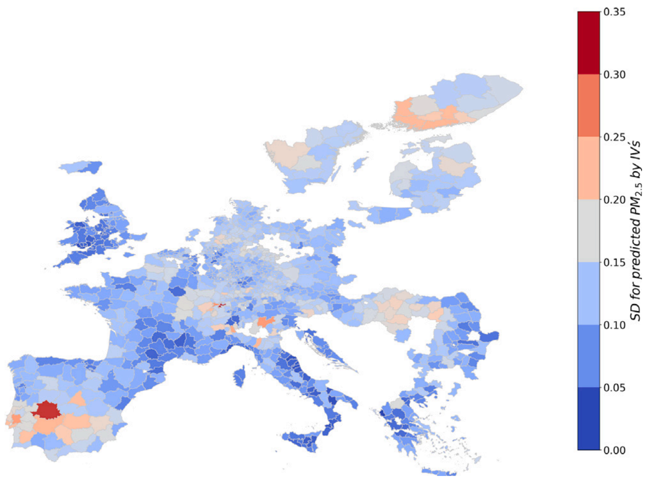

Fig. 5. Variation in predicted $PM_{2.5}$ concentration by NUTS-3 region.
Notes: Distribution over NUTS-3 regions of $PM_{2.5}$ induced by the two instruments in the estimation sample. It results from the predicted values of the control function (the 1st stage regression of the baseline model with both instruments) when setting all control variables at their mean. Colors reflect the standard deviation of predicted $PM_{2.5}$ concentration within the region over time.

projects are more valuable, the mix of short- and long-term projects in a given year would lead to an underestimation of innovation output. However, this is a more general issue, as it is well-established that patents vary substantially in value (Harhoff et al., 1999). To address this, we use several measures of quality-weighted innovation output as alternative dependent variable, finding that our effects become stronger when adjusted for quality.

4. Main results

4.1. Baseline

Table 2 presents the baseline results. The instrumental variables strongly affect air pollution. In line with our expectations, both instruments cause an increase in $PM_{2.5}$ concentration. Table 3 shows summary statistics of the predicted $PM_{2.5}$ concentrations resulting from our 1st-stage regressions. It examines variation in predicted $PM_{2.5}$ values induced only by the instruments when fixing controls at their mean. The last row shows that the standard deviation of fitted $PM_{2.5}$ concentrations is 0.38 $\mu g/m^3$ in our preferred specification (including both IVs in the control function). It ranges between 13.03 $\mu g/m^3$ and 15.69 $\mu g/m^3$. Fig. 5 plots the standard deviation of these fitted $PM_{2.5}$ concentrations across NUTS-3 regions. For the vast majority of NUTS-3 regions, the standard deviation of predicted pollution is above 0.05 $\mu g/m^3$, indicating the variation used for identification covers a large subset of our sample.

The first column of Table 2 shows OLS estimates of a model identical to Eq. (2) where we do not control for the control function residuals. The next four columns show the second- and first-stage results when instrumenting $PM_{2.5}$ concentration with thermal inversions or ventilation capability, respectively. The last two columns show the second- and first-stage results when instrumenting $PM_{2.5}$ concentration with both instrumental variables. The common IV test statistics confirm the validity of the assumptions underlying our analyses. The Kleibergen–Paap F statistics are high for all IV regressions and above critical values for weak instruments formulated in Stock and Yogo (2002).[20]

In line with expectations, OLS estimates are smaller than estimates obtained with the IV regressions, suggesting that omitted variables and reverse causality lead to underestimation of the harmful effects of air pollution. Standard OLS suffers from endogeneity issues due to several unobserved confounders. For instance, technological opportunity or profits could affect innovation while co-varying with economic activity, the main driver of pollution. Omitting them leads to *underestimating* the damage done by air pollution. Although less obvious, reverse causality could further attenuate this bias if R&D activities contribute to air pollution. Such a mechanism would further understate the costs of low air quality.

Including only the thermal inversion instrument results in a coefficient of 0.048 (standard error is 0.030). The ventilation capability instrument results in a coefficient of 0.096 (standard error is 0.034). Our preferred specification with both IVs yields a coefficient of 0.072 (standard error is 0.022). The fact that the different instruments, while both strong, produce different results is not necessarily surprising. The variation induced by thermal inversions is not geographically or temporally overlapping as they are caused by quite different weather phenomena. The difference in results suggests there may be significant heterogeneity across regions (which we explore in Section 4.3).

[20] F-statistics for the three control function regressions are all above the 10% (Stock and Yogo, 2002) critical values.

Table 2
Baseline results: Poisson with control function.

	No IV	Inversions IV		Ventilation IV		Both IVs	
Dep. Variable:	Patents	Patents	$PM2.5$	Patents	$PM2.5$	Patents	$PM2.5$
$PM2.5$	−0.0043*	−0.048		−0.096***		−0.072***	
	(0.0025)	(0.030)		(0.034)		(0.022)	
$Share\ Inversions$			4.21***				4.07***
			(0.58)				(0.57)
$Share\ low\ VC$					279.1***		267.0***
					(45.3)		(45.2)
$CF\ residuals$		0.044		0.092***		0.068***	
		(0.030)		(0.034)		(0.022)	
NUTS-3 FE	✓	✓	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
KP F-stat.			52.57		38.01		41.46
Mean PM2.5	13.78	13.78		13.78		13.78	
Mean Patents	123.11	123.11		123.11		123.11	
N	11724	11724	11724	11724	11724	11724	11724

Notes: Baseline estimates of the effect of air pollution on innovation. Innovation is measured by the number of patent filings in the upcoming year. The first column shows the results for the Poisson regression with no control function. The second and third columns show the second- and first-stage results of a Poisson model and a linear control function with thermal inversions as an instrument, where the first-stage residuals are controlled for in the second-stage estimation. In the fourth and fifth columns, we use ventilation capability as an instrument. The last two columns show the results when using both instruments jointly. $Share\,of\,Inversions$ measures the share of days in a given year on which a thermal inversion occurs. $Share\,of\,low\,VC$ measures, for a given year, the average number of consecutive days with a low ventilation coefficient (divided by 365 days). We define a low ventilation coefficient as any daily average ventilation coefficient that is equal or below the $20th$ percentile of the distribution of daily average ventilation coefficients within a NUTS-3 region. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3
Variation in $PM_{2.5}$ concentration induced by the instruments.

	N	Mean	Std. Dev.	Min	50th	Max
Predicted $PM_{2.5}$ with TI	11724	13.78	0.39	13.15	13.75	15.24
Predicted $PM_{2.5}$ with VC	11724	13.78	0.10	13.59	13.76	15.10
Predicted $PM_{2.5}$ with both IV's	11724	13.78	0.38	13.03	13.75	15.69

Notes: Summary statistics of predicted $PM_{2.5}$ concentrations estimated in first-stage regressions of our baseline model. Predicted values are obtained after first-stage regressions, setting control variables to their mean.

In our preferred specification (second to last column), an increase in the $PM_{2.5}$ concentration with 1 $\mu g/m^2$ causes a reduction in the number of patent filings per NUTS-3 region of 7.2%. To interpret the magnitude of the effect, we consider the year-to-year average decrease in pollution in Europe, 0.17 $\mu g/m^2$ as a reference point. This reference increase is well within the bounds of the variation exploited, as it is smaller than the standard deviation induced by our instruments in the overall sample. Following this logic, we conclude that Europe's average improvement in air quality has caused a 1.2% increase in inventive output. For the average region, this corresponds to 1.51 patents. To gauge the overall magnitude of this estimate, we can treat Eq. (2) as the structural invention production function and predict patented inventions when air quality would have remained at 2001 levels. We find that the total number of patents in Europe between 2001 and 2012 would have been 1.94 percent lower if it were not for the overall improvement in air quality.

Is this a large effect? In Section 6, we try to gauge the economic costs implied by the effect and compare it to the effect on same-year GDP from Dechezleprêtre et al. (2019). Here, we compare our effects to those found in the evaluation of R&D support policies. For instance, Myers and Lanahan (2022) show that a $1M Department of Energy grant results in 0.55 additional patents, while Bronzini and Piselli (2016) find that a €250k subsidy yields one additional patent. Our effect implies a reduction of 1.51 patents for the average region. The effect of our reference increase in air quality is then equivalent to a $3M subsidy (Myers and Lanahan, 2022) or €400k (Bronzini and Piselli, 2016). This suggests that reducing air pollution could have an impact comparable to substantial R&D grants.

Patent quality. To investigate whether air pollution affects the quality of observed inventions, we re-estimate our baseline specification after weighting patent counts for the total number of citations patents receive, a common proxy for the quality or technological importance of patents (Trajtenberg, 1990; Hall et al., 2001). Next to total citation counts, we use two alternative measures for the number of citations. First, we count the number of citations patents received within a 5 year window to leave each invention the same amount of time to be cited. Second, we normalize the total number of citations received by how often the average patent in the same year and technological field is cited. Such normalization addresses the concern that citation practices differ across

Table 4
Weighting for patent quality.

Dep. Variable	Patents	Cites	5-y cites	Norm. cites	Fam. size	Triadic	EPO filing	EPO grant	Any grant
$PM2.5$	−0.072***	−0.12***	−0.11***	−0.094***	−0.11***	−0.090**	−0.091***	−0.070***	−0.070***
	(0.022)	(0.036)	(0.038)	(0.030)	(0.032)	(0.043)	(0.024)	(0.026)	(0.023)
CF $residuals$	0.068***	0.12***	0.11***	0.096***	0.11***	0.093**	0.088***	0.068***	0.070***
	(0.022)	(0.036)	(0.038)	(0.030)	(0.031)	(0.043)	(0.023)	(0.025)	(0.023)
NUTS-3 FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
KP F-stat.	41.46	40.55	39.71	40.55	41.46	37.09	39.99	40.19	40.18
Mean PM2.5	13.78	13.78	13.75	13.78	13.78	13.36	13.75	13.56	13.75
Mean Dependent	123.11	594.84	201.32	44.27	585.50	18.96	60.23	33.92	76.44
N	11724	11688	11592	11688	11724	10344	11640	11160	11652

Notes: Estimates from the second-stage Poisson regression with quality-weighted patent counts as dependent variable. The first column repeats the baseline result, where *Patents* is the number of inventions for which a patent was filed in a region. *Cites* is the total number of citations received by all patents in a region. *5-y cites* only counts citations within a 5 year window from the patent filing. *Norm. cites* normalizes the number of citations of a patent by how often patents in the same filing year and technological field are cited on average. *Fam. size* weights the number of inventions by how many distinct filings it was the subject of. *Triadic* counts only inventions for which a patent was filed in the major patent jurisdictions (USPTO, EPO, and JPO). *EPO filing* counts only inventions for which a patent was filed at the EPO. *EPO grant* counts only inventions for which a patent was granted at the EPO. *Any grant* counts only inventions for which any patent was granted (at the EPO or elsewhere). All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

technologies and avoids having to specify a time window in which citations received are counted. Each measure is summed up within NUTS-3 region and year to obtain 'citation-weighted' patent counts. In addition, we employ two measures that proxy the private value of inventions by looking at how widely patent protections is sought by the applicant. The idea is that, because patenting is costly, only valuable inventions will merit the filing of a large number of individual patents (Harhoff et al., 2003). To implement this, we sum up the number of individual filings related to any given invention in our set to obtain a 'family-size-weighted' patent count. Alternatively, we count only inventions with a so-called 'triadic patent', which means it has a filing in the US, European, and Japanese patent offices (Aghion et al., 2016), and we count only inventions filed at the EPO, meaning the applicant intended to file in at least 2 European countries. Finally, we use the grant status of a patents a signal of quality. Here we restrict our patent count variable to patent families with at least 1 EPO grant and to those with at least one grant anywhere.

Table 4 presents the result when implementing the baseline specification with each of these quality-weighted measures as the outcome variable. Across the board, weighting for patent quality never makes the negative effect of air pollution significantly weaker. Coefficients are between 0.070 (2.7% lower than the baseline) and 0.12 (67% higher than the baseline). Table 10 in Appendix A.1 explores the effect of air pollution across the quality distribution. For this analysis, we assign each patent to its quality quartile (using citations and family size as quality measure) and run the baseline specification with patent counts in each quartile as the outcome variable. For both quality measures, the upper quartile effect is larger than the lower quartile effect, confirming the findings for quality-weighted patent counts.[21]

4.2. Robustness checks

The remainder of this section discusses various robustness tests of our baseline result. Fig. 6 presents an overview of the coefficient of interest for each of these tests. Appendix A.1 shows the detailed results.

Placebo test. We conduct a simple placebo test in which we estimate the effect of air pollution in year $t + 1$ on innovation output in year t.[22] This test is useful in that finding such an effect would suggest that factors other than pollution exposure create a correlation between the instruments and innovation. Reassuringly, we find small and statistically insignificant coefficients (Table 11).

Timing of the effect. We examine the timing of the effect by lagging and leading innovation output. Table 12 reports the coefficients in our baseline specification when innovation is measured using patents in year t, $t + 1$ (the baseline specification), $t + 2$, $t + 3$, ad $t + 4$. The effect of air pollution on innovation only appears for our baseline specification. We further examine the timing of the effects by splitting up our patent counts variable by the quarter in the year the patent is filed for. Fig. 8 confirms the conclusion that effects are most pronounced in the year following the air pollution exposure. These patterns lend credence our assumption about the lag between innovation activity and patent filing. Furthermore, it suggests that displacement in time of innovation (where innovation not performed in polluted years is carried out in later years) is unlikely.

[21] The middle quartile results show either a much larger or much smaller coefficient than the outer quartiles. It is not clear why this is the case, but taking the middle quartiles together, the result is close to the baseline coefficient.

[22] Note that innovation output is still measured in the year after the innovation presumably occurred. All controls are measured in year t as in the baseline specification.

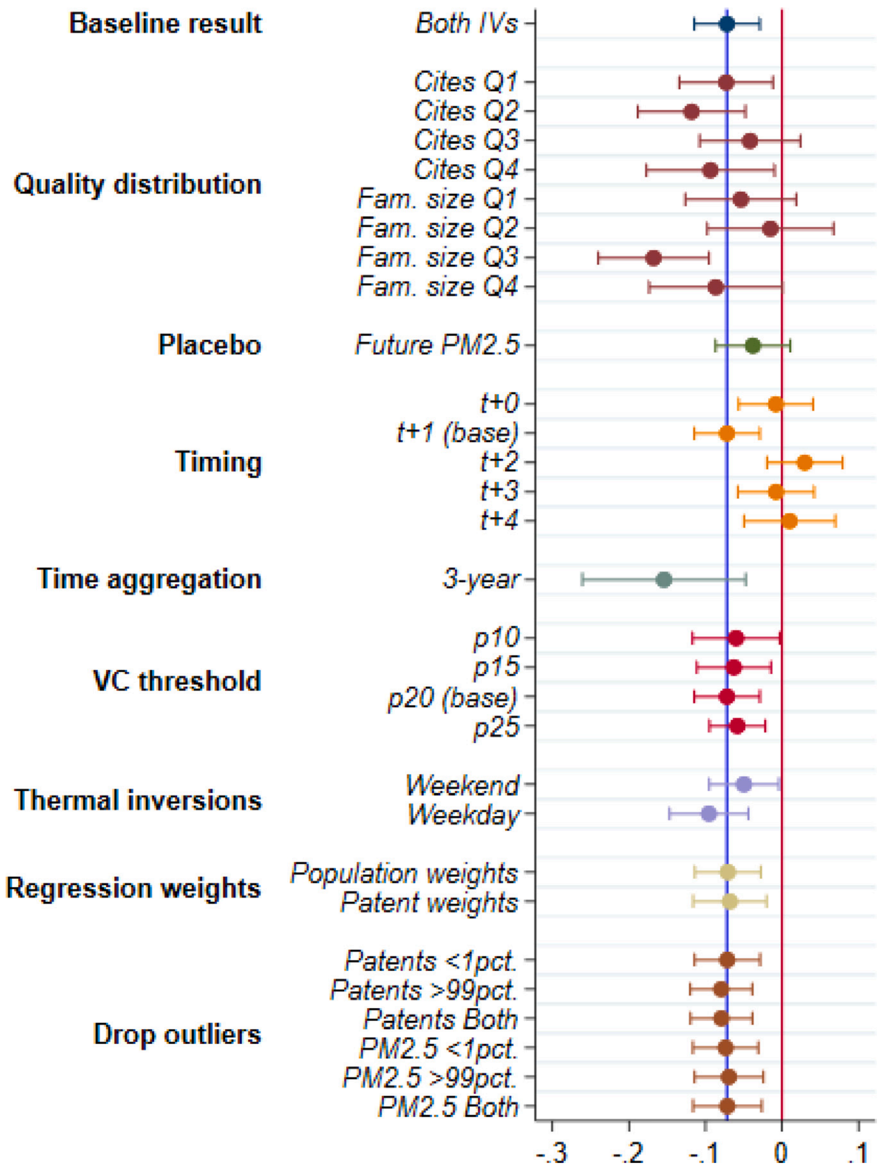

Fig. 6. Coefficient plot of robustness results.
Notes: Summary of the robustness checks described in Section 4. Plots the coefficient of predicted $PM_{2.5}$ in the second stage regression, with 95% confidence intervals. **Baseline result** refers to the baseline specification with both IVs. **Quality distribution** refers to specifications that count patents in each quartile of the quality distribution. **Placebo** refers to the regression estimating the effect of future air pollution on current innovation output. **Timing** refers to specifications with patent output in time periods as outcome variable. **Time aggregation** refers to specifications in which the unit of time is aggregated. **VC Threshold** refers to regressions with an alternative definition of the ventilation capability instrument. For *p10*, *p15*, *p20*, and *p25* we use the definition of the ventilation coefficient as low when it is equal or below the respectively 10th, 15th, 20th, 25th percentile of the distribution of daily average ventilation coefficients within a NUTS-3 region. **Thermal inventions** split up the instrument by weekend and weekday inversions. **Regression weights** applies population or patent weights when estimating the baseline regression model. **Drop outliers** refers to regressions in which we drop regions that are outliers in terms of patents or air pollution.

Alternative time aggregation. An alternative approach to test sensitivity to our timing assumption is to aggregate the data in wider time windows. The downside is that IV-induced variation decreases as the weather conditions are ‘averaged out’ when looking at longer time periods. Table 13 shows the results when using 3 year, 4 year, and 6 year windows to aggregate our data along the panel dimension. Patenting output is not lagged for this exercise. For 3 years windows, the coefficient becomes −0.15 – about double the baseline coefficient. However, the KP F-statistic drops to 13.5, which is below established thresholds for strong instruments. With this caveat in mind, it is reassuring (if anything) that the effect size increases, suggesting our baseline approach is on the conservative side. Further aggregating to 4 year windows and 6 year windows results in KP F-statistics below 4, making clear that our design is not valid for such analyses.

Table 5
Heterogeneity sector.

Sector	All	Chemistry	Elec. Engin.	Instruments	Mech. Engin.	Other
$PM2.5$	−0.072***	−0.11**	−0.072**	−0.049	−0.067**	−0.058
	(0.022)	(0.044)	(0.035)	(0.040)	(0.033)	(0.046)
$CF\ residuals$	0.068***	0.10**	0.069**	0.044	0.067**	0.057
	(0.022)	(0.044)	(0.035)	(0.039)	(0.033)	(0.047)
NUTS-3 FE	✓	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
KP F-stat.	41.46	35.96	43.67	33.79	40.98	33.62
Mean PM2.5	13.78	13.42	13.71	13.35	13.72	13.39
Mean Patents	123.11	23.90	30.87	18.51	45.34	10.90
N	11724	10920	11184	10656	11376	10656

Notes: Estimates from the second-stage Poisson regression with the number of patented inventions, split up by technological sector, as dependent variable. Patents are assigned to sectors based on the taxonomy developed in Schmoch (2008). The first column repeats the baseline estimates. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Ventilation Capability thresholds. We check the sensitivity of the results to the way in which we construct our ventilation capability instrument by altering the threshold for which we define a ventilation coefficient to be 'low'. The baseline specification uses the 20th percentile of the distribution of daily average ventilation coefficients in a region. Table 14 shows that our main result holds when using the 10th, 15th, 20th, and 25th percentile as a threshold.

Weekend vs. weekday inversions. We borrow from Avila Uribe (2023) to do an interesting robustness check to our thermal inversions instrument. The idea is that output of (knowledge) workers should be mainly affected by high air pollution concentrations on working days. We can therefore check the design by differentiating between weekend day and weekday inversions. Table 15 shows that, when using just the thermal inversions IV, we find a negative effect for weekday inversions only. When adding our second IV (calculated as in the baseline), we see that the baseline effect goes from −0.072 to −0.095.[23] This again suggests our baseline approach errs on the conservative side.

Regression weights. The studies on air pollution and GDP present coefficients from regressions with population weights to ensure that effects in smaller areas do not drive the main result (Dechezleprêtre et al., 2019; Avila Uribe, 2023). In our setting, this seems less of a problem because smaller regions will also have lower patent intensity, so that when interpreting the result in terms of number of patents lost due to air pollution, such effect should be accounted for. This is not the case when looking at GDP growth or GDP per capita. Nevertheless, it is an interesting check to see if the results hold when granting more weight to larger and more patent-intensive areas. We use the first year of the panel to construct weights base on logged population and logged number of patents, and re-run our baseline regression.[24] As can be seen in Table 16, the results remain very close to the baseline results.

Dropping outliers. We verify whether our results hold when dropping regions that are outliers in terms of patenting or air pollution. We drop the bottom and/or the top percentile of regions in terms of total patents or average air pollution. Table 17 shows that results are robust.

4.3. Heterogeneity regions

Understanding effect heterogeneity across regions can help inform policymakers about where abatement efforts are best targeted. We explore whether our result varies by technological sector, by patent intensity, by pollution level, and by level of urbanization.

By technology sector. Table 5 explores sectoral heterogeneity by assigning patents to one of 5 broad technological categories as defined by Schmoch (2008), who assigns IPC codes to 35 technology fields nested into these sectors: Chemistry, Electrical Engineering, Instruments, Mechanical Engineering, and Others. The coefficients vary (slightly) by sector. Electrical Engineering and Mechanical Engineering show effects close to the baseline. In Chemistry the effect is about 50% larger, while in Instruments and Other it is 32% and 19% smaller (and not statistically significant at the 10% confidence level). It should be noted that these sector-specific estimates are less precise than the baseline effect, which is within 1 standard error for each sector. As a result, we conclude that there is no evidence for (strong) heterogeneity by sector.

This seemingly contradicts the findings of heterogeneity by sector when looking at GDP (Dechezleprêtre et al., 2019; Avila Uribe, 2023). However, GDP captures very heterogeneous activities, which may not be equally sensitive to poor air quality. Our setting

[23] These results hold when restricting to granted patents only, and the effect for weekday inversions is present across the quality quartiles as discussed above. The relevant tables are not reported and are available upon request.

[24] Not doing the log-transformation puts extremely large weights on a few big regions and make our IVs weak.

Table 6
Heterogeneity by patenting output and level of pollution.

	Patents		$PM2.5$	
Sub-sample	Below median	Above median	Below median	Above median
$PM2.5$	−0.16**	−0.063***	−0.016	−0.10***
	(0.071)	(0.021)	(0.030)	(0.031)
CF *residuals*	0.17**	0.059***	0.024	0.096***
	(0.070)	(0.021)	(0.032)	(0.030)
NUTS-3 FE	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
KP F-stat.	15.26	26.63	30.31	18.39
Mean PM2.5	14.37	13.30	10.89	16.95
Mean Patents	12.93	215.87	104.55	143.65
N	5292	6360	6072	5580

Notes: Split-sample estimates from the second-stage Poisson regression with the number of patented inventions as the dependent variable. The first two columns split the sample by the median number of patents in a region (where patents are averaged across time for each region). The second two columns split the sample by the median $PM2.5$ in a region ($PM2.5$ is averaged across time for each region). All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

focuses on (high-performance) cognitive activity, and we should expect less heterogeneity as the type of work performed is relatively similar across workers.

By Patent Intensity and Pollution Levels. Table 6 divides the sample into regions above and below the median levels of total patents and average air pollution. The effect size is larger but less precisely estimated for low-patenting regions (0.16 in the below-median versus 0.063 in the above-median sample). When examining the number of patents gained with an improvement in air quality, the economic effect is significantly greater in patent-intensive regions. For our reference increase in air quality, we would see 0.42 more patents in low-patenting regions and 2.31 more patents in high-patenting regions. The right-hand side panel of Table 6 shows the sample split by air pollution levels. The coefficient of interest is small and insignificant in less polluted regions, whereas in more polluted regions, the effects are large and statistically significant. For these regions, a 1 $\mu g/m^2$ increase in the $PM_{2.5}$ concentration reduces the annual number of patent filings by 10%. Being driven by high-pollution regions, the effect of pollution on innovation appears to be non-linear.

Urbanization level. Table 7 splits the sample into regions that are (1) predominantly rural, (2) intermediate, and (3) predominantly urban – based on Eurostat's 'Rural Development' indicators.[25] The effect size for Urban is 28% above the baseline. For Rural it is also above the baseline but very imprecisely estimated. For Intermediate, we have a precisely estimated effect that is 21% below the baseline effect. Given these results, we should be cautious about drawing strong conclusions regarding heterogeneity, as the strength of the instruments decreases significantly in the split-sample analysis. Nevertheless, we can examine the potential increase in the number of patents resulting from a decrease in air pollution of 0.17 $\mu g/m^2$. For urban areas, this reduction would result in an increase of 4.27 patents; for intermediate areas, 1.02 patents; and for rural areas, 0.61 patents. This suggests some degree of heterogeneity, though the estimates come with uncertainty.[26]

Taken together, the results are consistent with an interpretation where the effect size increases with pollution levels (or where a threshold effect exists, with negative effects above certain pollution levels). The economic implications, specifically the abatement benefits, depend heavily on the amount of innovative activity in a region because the heterogeneity in effects is much smaller than the heterogeneity in patenting activity. If we assume abatement costs are similar across regions, the most efficient policy would target abatement efforts in urbanized regions, which have many patents and relatively high pollution.

5. Mechanisms

This section provides evidence for our assertion that we can translate the region-level estimates to economy-wide costs which are not captured in previous estimates. To this end, we address three concerns. First, we examine whether inventor mobility between regions can explain the results. If air pollution causes inventors to relocate, at least a portion of the human capital loss in one region is recovered by the human capital gain in another one. If this mechanism is important, aggregating our estimates will lead to overestimating the economic costs induced by innovation. Second, we discuss to which extent shocks to R&D funding might drive

[25] For more information, refer to https://ec.europa.eu/eurostat/web/rural-development/methodology.

[26] These results hold when restricting to granted patents only, and the effect in urbanized regions manifests itself across the quality quartiles as discussed in Section 4 4.2. The relevant tables are not reported and are available upon request.

Table 7
Heterogeneity by patenting output and level of urbanization.

Sub-sample	Urban	Intermediate	Rural
$PM2.5$	−0.092**	−0.057**	−0.084
	(0.038)	(0.028)	(0.053)
$CF\ residuals$	0.085**	0.054**	0.090
	(0.037)	(0.028)	(0.055)
NUTS-3 FE	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓
Country-Year FE	✓	✓	✓
Controls	✓	✓	✓
KP F-stat.	11.60	15.09	17.76
Mean PM2.5	14.49	14.02	12.85
Mean Patents	272.78	105.55	42.49
N	2616	5292	3612

Notes: Split-sample estimates from the second-stage Poisson regression with the number of patented inventions as the dependent variable. The columns show the results for the sub-sample of NUTS-3 regions classified as respectively predominantly urban (first column), intermediate (second column), and predominantly rural (classification is made according to Eurostat's 'Rural Development Indicators'). All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8
Effect on inventor migration.

	Emigration		Immigration		Both
Dep. Var.	Inventors	Inv. patents	Inventors	Inv. patents	Pat. per cap.
$PM2.5$	−0.029	−0.096	0.0058	−0.091	−0.067***
	(0.094)	(0.16)	(0.080)	(0.13)	(0.025)
$CF\ residuals$	0.018	0.090	0.0094	0.10	0.060**
	(0.093)	(0.16)	(0.081)	(0.13)	(0.025)
NUTS-3 FE	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
KP F-stat.	34.90	34.87	37.35	36.47	41.46
Mean PM2.5	13.23	13.24	13.14	13.15	13.78
Mean Dep. Var.	2.55	10.04	2.72	10.54	0.00
N	9360	9300	8976	8940	11724

Notes: Second-stage results of the analyses of the effect on inventor migration. *Inventors* refers to the number of inventors that were identified as moving between regions. *Inv. patents* weights the number of migrating inventors by the number of patents inventors filed within 5 years before the relocation. The first two columns refer to inventors migrating out of the focal region, the next two columns refer to inventors migrating into the focal region. The final column estimates the effect on the number of patents in a region divided by the region's population. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

our results. While this mechanism does not preclude us from aggregating across regions, it would invalidate our claim that the implied costs are not captured by the effect on short-term economic activity (Dechezleprêtre et al., 2019). Indeed, reductions in R&D spending are captured by measures of GDP. Finally, we examine the effects of air pollution on total mortality in our sample. Such increases in mortality would be picked up by previous estimates of the economic costs of air pollution as assessed by losses in productive life years (Landrigan et al., 2018).

5.1. Inventor mobility

To identify relocation events and their timing, we chronologically order all inventors' addresses stated on their patents and mark switches between NUTS-3 regions. The relocation year of an inventor can be easily determined when such a switch happens within the same year or within two consecutive years. In this case, we assume the relocation year corresponds to the year preceding the filing date of the latest patent stating the inventor's new address.[27] When a switch co-occurs with a gap of multiple years, we use the median as the relocation year.[28] To exclude unreasonable movement patterns, we conduct the data-cleaning process proposed by Zacchia (2018) (see Appendix A.2). Using these relocation events, we construct our measure of inventor relocation by counting the yearly number of inventor movements out of and into NUTS-3 regions for the sample period 2001–2012. To verify whether our decision rule to select the relocation event year drives the results, we re-estimate our regressions using a sample with only relocation events that can be precisely determined. To account for the fact that relocation of more prolific inventors has a relatively large impact on a region's innovation output, we construct an additional measure in which we weigh each relocation event by the number of patents respective inventors filed within five years before the relocation year.

It is important to mention that we can only observe the migration of inventors that file patents before and after a relocation event. Migration before the first filed patent and after the last patent filing, as well as multiple moves between two filings, cannot be tracked. As such, we have data for a sub-sample of relatively prolific relocating inventors. While we cannot fully address the resulting selection issue, we note resulting bias would plausibly lead us to overestimate how much of our effect can be explained by migration. Indeed, the number of inventions produced in the region receiving the migrating inventor in response to increased pollution is larger for those more productive individuals identified by our method.

Overall, 1,295,300 inventors who filed patents between 1944 and 2013 state addresses in NUTS-3 regions. Only 5.38% of those inventors file patents in two different regions, suggesting relocation is a rare event. Fig. 9 in Appendix A.1 shows relocation events of inventors between NUTS-3 regions for the period 2001–2012. Similarly, Fig. 10 shows relocation events of inventors between NUTS-3 regions and regions outside Europe. Unsurprisingly, most relocation events occur for areas with many patent filings.

Table 8 shows the results of applying our baseline model to estimate the impact of air pollution on the annual number of inventor emigration events (first pair of columns), immigration events (second pair of columns), and the number of patents per inhabitant of a region. The latter measure's logic is that if the effects on migration flows are important relative to the amount of inventive output, we should see that the output per capita is considerably different from our baseline effect. Both for immigration and for emigration, the results show no statistically or economically significant results. To verify that the estimated coefficients cannot explain our baseline results, consider the upper bound of the 95% confidence interval around the emigration coefficient (0.16). Taking this as the true effect would imply that inventor migration in response to a 1 $\mu g/m^2$ increase in $PM_{2.5}$ would increase with 16%. The baseline migration rate is about 2.5 inventors per year, which implies an outflow of inventors with less than half of an inventor. With about 0.8 patents per year per migrating inventor, this results in a reduction of 0.4 patents due to migration. The same increase in air pollution reduces the average number of patents in a region by about 9. As such, it becomes clear that migration could only marginally account for our baseline result, even for this highly conservative calculation. This pattern is further confirmed by the patents per capita analysis, which yields a coefficient of 0.067, about 7% smaller than the baseline coefficient. Table 18 in the appendix presents the migration regressions for the high-confidence relocation events for which two different locations occur on patents within the same year or within two consecutive years. Again, IV estimates show small and insignificant coefficients. In conclusion, we do not find evidence that the migration channel can go a long way in explaining our baseline effects. These patterns are consistent with the idea that variation in air pollution induced by atmospheric conditions is too idiosyncratic and small to cause inventors to move into or out of a region.[29]

5.2. R&D expenditures

While we do not observe R&D expenditures for enough firms and regions to test this mechanism directly, we provide two pieces of evidence making the R&D spending story unlikely.

First, it is implausible that air pollution would significantly affect R&D investment within the short term of a year. Firms typically plan their R&D budgets well in advance, and short-term fluctuations in air quality are unlikely to prompt immediate adjustments in these plans. Given the year-to-year variability in air pollution driven by weather conditions, it is unlikely that firms would respond by altering their R&D expenditures. Consequently, the observed impact on innovation is more likely to be direct rather than mediated through changes in R&D spending.

Second, our findings in Section 4.1 indicate that quality-weighted patent counts are more affected by air pollution, which contradicts the notion that the mechanism is reduced R&D spending. If firms were cutting back on R&D, we would expect them to first eliminate lower-quality projects. However, the disproportionate reduction in high-quality patents suggests that the impact of air pollution directly influences the productivity of innovative activities rather than merely reducing the financial resources allocated to R&D. This inconsistency with the R&D spending mechanism further supports our claim that the effects we observe are not primarily driven by changes in R&D investment.

[27] We take the preceding year to account for the 3–9 months lag between invention and patent application.

[28] Decimal numbers are rounded up assuming the relocation year is closer to the patent filed under the new address.

[29] Clearly, our analysis does not rule out that inventors move within NUTS-3 regions. Our baseline estimates account for lost output implied by such moves.

Table 9
The effect of air pollution on overall mortality.

Dep. Var.	Mortality		Mortality rate	
	1st stage	2nd stage	1st stage	2nd stage
$PM2.5$		0.0053*		0.0073**
		(0.0031)		(0.0034)
$Share\ Inversions$	4.19***		4.19***	
	(0.57)		(0.57)	
$Share\ low\ VC$	271.3***		271.1***	
	(45.1)		(45.0)	
NUTS-3 FE		✓		✓
NUTS-3 Slopes		✓		✓
Country-Year FE		✓		✓
Controls		✓		✓
KP F-stat.		43.5		43.6
Hansen J stat. P-Val		0.62		0.53
Mean PM2.5		13.80		13.80
Mean Dep. Var.		3650.76		1.04
N	11688	11688	11688	11688

Notes: Estimates of the effect of air pollution on the natural logarithm of the total number of deaths (mortality) and the death rate (mortality rate) by NUTS-3 region. The table shows the results from 2SLS regressions with two instruments. $Share\,of\,Inversions$ measures the share of days in a given year on which a thermal inversion occurs. $Share\,of\,low\,VC$ measures, for a given year, the average number of consecutive days with a low ventilation coefficient (divided by 365 days). We define a low ventilation coefficient as any daily average ventilation coefficient that is equal or below the $20th$ percentile of the distribution of daily average ventilation coefficients within a NUTS-3 region. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The mortality estimates but not the mortality rate estimates control for population (in logs). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3. Mortality

Here we investigate whether the reduction in inventive output can be explained by an increase in mortality induced by air pollution. We apply the regression model described in Section 3.3 and estimate the immediate effects of air pollution on the annual number of deaths and annual death rate per NUTS-3 region. As we do not have reliable estimates of the number of inventor deaths, we conduct this analysis using total deaths per NUTS-3 region obtained from Eurostat. Table 9 shows the impact of $PM_{2.5}$ concentration on the natural logarithm of the total number of deaths per NUTS-3 region and the number of deaths per capita. The IV regressions show small and (marginally) significant coefficients. It is unlikely that mortality can explain our baseline effect. Assume the true effect is the upper bound of the 95 percent confidence interval in column 2 (0.011). In this case, our reference increase in air pollution of 0.17 $\mu g/m^3$ leads to 7.1 deaths in the average region. Knowing that the number of people who have filed for a patent in the last 10 year for the average region is 577, while the average population is 371,429, we get that about 0.15% of the population is an inventor. As a result, the mortality mechanism would account for the death of 0.011 inventors.

These small effects on mortality in our research design are consistent with recent quasi-experimental evidence of the mortality impact of air pollution. Using wind patterns around LA highways to instrument for air pollution, Anderson (2020) finds an increase in three-year mortality rates in people aged over 75, but no effect in age groups 55–64 and 65–74. With 91% of Europe's population aged below 75, it comes as no surprise we detect little of an effect on population mortality.[30] Using variation in pollution reduction thanks to the Acid Rain Program, Barreca et al. (2021) finds a reduced mortality on the overall population only after a period of three years of exposure, which is in line with the fact that we find small effects in our in one-year exposure windows.

These findings suggest that mortality cannot explain the reduction in inventive output from our baseline estimates as long as pollution-induced mortality for inventors is similar to that in the general population. Two arguments justify this assumption. First, inventors appear to have an age profile similar to that of the general population. Using estimates from Kaltenberg et al. (2021) of the birth year of inventors that file for a USPTO patent in 2011, we find that the median inventor is aged 44, as compared to a median of 41.6 for the EU population.[31] Furthermore, very few inventors are in the age category for which immediate mortality effects have been found (Deryugina et al., 2019; Anderson, 2020), with less than 5 percent of them older than 65, and less than 1 percent above age 75. Second, we see no reason why inventor mortality would be more responsive to air pollution than the general population.

[30] Based on population estimates of the United Nations, see https://population.un.org/wpp/.
[31] Based on Eurostatdata.

Prior research shows that inventors have high incomes, high levels of education, and high-income parental backgrounds (Bell et al., 2019; Aghion et al., 2017). These socioeconomic factors negatively correlate with mortality (Torssander and Erikson, 2009), possibly because of more healthy lifestyles and access to high-quality healthcare. If anything, we would expect lower responsiveness, making our estimates here err on the conservative side.

6. Cost calculation

This section describes a back-of-the-envelope calculation of the costs implied by the effect of reduced air quality on innovation. The baseline estimates suggest that our reference increase in air quality of 0.17 $\mu g/m^3$ in a given year costs the average region 1.51 patents. To translate this loss into economic value, we rely on attempts in the literature to value the average patented invention. Multiplying our baseline effect with estimates of the private and spillover value of innovative output gives the 'hidden' cost of air pollution under two assumptions. First, the innovations lost due to air pollution have a value similar to the average one. Second, the estimated effect is not picked up in prior cost estimates. The previous sections contain evidence supporting these assumptions.

This cost calculation heavily relies on estimates of the average private returns a patented invention generates over its lifetime. The innovation literature has adopted three approaches[32] to value patented inventions.[33] A first approach uses data on 'renewal fees' that a patent holder must pay to keep the patent alive. The decision to pay such fees is modeled to parameterize the private value distribution of underlying invention (Pakes, 1986; Schankerman and Pakes, 1986). A second approach uses the market valuation of firms in response to patenting, either using Tobin's Q regressions on patent stocks in yearly panels or using event studies of patent grants (Hall et al., 2005; Bessen, 2009; Kogan et al., 2017). A final approach surveys inventors from a representative sample of patents and inquires about their strike price for selling the IP right associated with the invention (Harhoff et al., 1999; Gambardella et al., 2008). Each of these approaches has been implemented for various samples of patents that differ in the time window and geographic scope considered.

We select three estimates from the literature – one for each approach – that are obtained using a sample most comparable to ours. Using renewal fee payments, Deng (2007) finds an average private value of 2.4 million euros[34] for granted EPO patents applied for between 1978 and 1996. Using Tobin's Q regressions, Bessen (2009) finds an average value of 413,000 euros in a sample of USPTO patents issued between 1979 and 1997.[35] Finally, Gambardella et al. (2008) find an average value of 3.5 million euros using 8217 responses from surveys of inventors listed on EPO patents filed between 1993 and 1997. We believe these estimates provide a reasonable range for our sample's private value of inventions.

The value of spillovers generated by patented inventions is a second input to our cost calculation. The literature typically estimates the *ratio* of spillover returns and private returns. Using our private value estimates above, we can use these estimated ratios to obtain the value of spillovers generated by the average patent in our sample. Under the assumption that spillovers make R&D of others more productive, their value can be estimated by regressing the innovation output of a certain firm on the R&D performed by other firms. The main challenge is to define the pool of spillover receivers and properly instrument for endogenous R&D spending. Exact magnitudes depend on how these challenges are addressed, but the value of spillovers is generally found to be at least as large as the private value. Bloom et al. (2013) estimate the marginal private returns to R&D to be 21%, and the marginal spillover returns to be 55%, a ratio of 2.62. Focusing on spillovers that emanate through personal connections of inventors, Zacchia (2020) finds a ratio of 1.17. Accounting for the fact that some spillovers are internalized through the market for technology, Arque-Castells and Spulber (2022) find a ratio of 5.29. Myers and Lanahan (2022) adopt a slightly different approach and estimate the effect of windfall R&D subsidies to firms on their own and others' innovation output. They find a spillover-to-private-value ratio of nearly 3. For our back-of-the-envelope calculation, we multiply these ratios with various private value estimates detailed above.

Based on these back-of-the-envelope calculations, Fig. 7 plots the cost implied by our reference air pollution increase as a share of the average region's GDP. The x-axis displays a range of private value estimates between 400,000 and 5 million euros. The vertical black lines show the three private value estimates we use from the literature described above. The horizontal black line indicates the estimate in Dechezleprêtre et al. (2019) (DRS) to compare the innovation and same-year GDP effect of air pollution. The colored lines plot the estimated costs for different assumptions about the relative spillover returns – one where spillovers are assumed not to exist and three estimates from the literature described above.

The implied costs of air pollution are high and – even for rather conservative scenarios – comparable to the effect of air pollution on same-year GDP. The most conservative method assumes that inventions have an average value of 413,000 euros (creating no spillovers). This scenario implies a cost of 0.0076% of the GDP for our reference increase in fine particulate matter. For the average region, this amounts to 0.62 million euros, about 5.7% of the estimate in DRS. If we would take the survey-based private value estimate of 3.5 million euros and the high spillover ratio found by Arque-Castells and Spulber (2022), our effect implies a cost of 0.41% of the GDP. This is about triple the effect found in DRS.

[32] A fourth approach uses the price of traded patents as the upper/lower bound of the value of an IP right on the underlying invention, as assessed by their seller/buyer. For instance, Fischer and Leidinger (2014) collect a dataset of 617 patents filed at the United States Patent and Trademark Office (USPTO) to correlate patent characteristics to the price of sold patents. While this approach is superior in precision, the main disadvantage for our purposes is the limited availability of data on traded patents and the fact that resulting samples are unlikely representative of the population.

[33] For an excellent discussion, see Stevenson et al. (2022).

[34] We express all values in this cost calculation in EUR 2010

[35] Using an event study approach, Kogan et al. (2017) find a value of 17.9 million in a sample of USPTO patents issued between 1926 and 2010. We do not consider this estimate in our cost calculation as it seems unrealistically large – 5 to 10 times larger than other estimates – in the context of our sample of EPO patents.

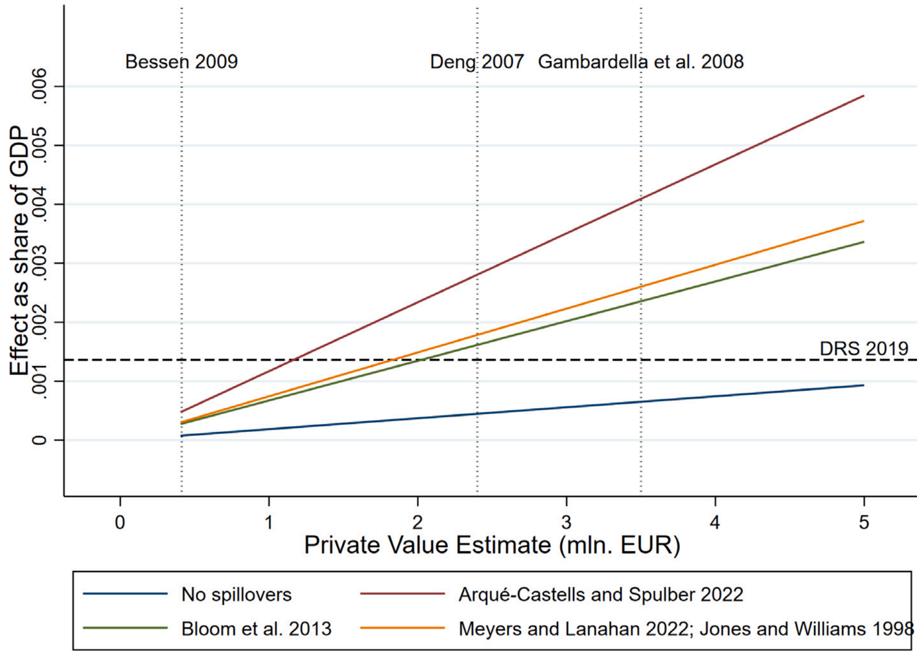

Fig. 7. Costs as share of GDP.
Notes: Results of cost calculation exercise. The y-axis shows the cost as a share of GDP implied by the baseline effect of our reference increase in $PM_{2.5}$ concentration of 0.17 $\mu g/m^3$. It converts the average effect of 1.51 patents into economic costs by multiplying it by the average value of a patent, and dividing it by the average region's GDP. The implied cost is displayed as a function of the private value estimate (x-axis). The colored lines reflect different estimates of the value of spillovers. The blue line assumes no spillover. The green, yellow, and red lines assume spillover ratios estimated in the literature of respectively 2.62, 3, and 5.29. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Clearly this is a very wide range. Our preferred estimate takes the most conservative estimate that is based on European patents (Deng, 2007). However, this estimate is based on EPO granted patents only, while only 27.6% of inventions in our sample belong to that group. We adjust Deng's value by comparing the citation rates of non-granted EPO patents (4.3 citations) and EPO granted patents (8.3 citations). Assuming this ratio accurately reflects the value differential, we apply it to modify Deng's estimate (2.4 million) for our sample and arrive at an adjusted value of 1.6 million. For spillover effects, we use the most conservative estimate from Bloom et al. (2013), which considers the full spillover effects without being based on one specific sector. With these assumptions, the implied economic cost 0.11 percent of the GDP, which is about 78% of the effect found in DRS.

7. Conclusion

Recent findings suggest that improving air quality can considerably advance immediate economic output. We argue that air pollution may also cause harm to future growth by reducing innovation output. To examine the importance of this mechanism, we estimate the effect of air pollution on innovation using two weather phenomena as sources of exogenous variation.

We find a large negative effect of air pollution on innovation. The estimates suggest that a decrease in pollution of 0.17 $\mu g/m^3$ – the yearly drop observed in Europe – leads to 1.2% more patented innovations. We find no evidence for a concurrent decline in average patent quality, so we interpret our coefficients as effects on total innovative output. We also find that the negative effect of air pollution is more pronounced in more polluted regions, and that patent-intensive regions lose more innovation output in absolute terms. Targeting urbanized regions with abatement efforts would create the largest benefits because they are relatively patent-intensive and relatively more polluted.

Two additional findings suggest that our estimates reflect costs not captured in prior assessments. First, the results cannot be explained by inventor migration. As such, negative effects on innovation in one region are unlikely to be made up for by more innovation in a different region through the movement of human capital. Given the absence of such spillover effects, we believe our results can be aggregated to the national and international levels. Second, shocks to R&D budgets or increased mortality induced by air pollution are unlikely to explain the results. These findings suggest that the economic costs implied by reduced innovation output are not captured in previous assessments based on reduced human capital and short-term economic output (Landrigan et al., 2018; Dechezleprêtre et al., 2019).

A back-of-the-envelope calculation shows that the cost implied by the effect of air pollution on innovation is large. Relying on the literature for private and spillover value estimates, we calculate that a reasonable improvement in air quality can create a value of 0.11% of the GDP for the average region. Compared to prior estimates, this suggests that the economic costs of air pollution increase by three quarters when accounting for the effect on innovation.

While we see our findings as relevant to inform cost–benefit analyses leading to abatement strategies, it is important to stress three limitations of our study. First, we cannot affirm the importance of different potential mechanisms behind our results. The patent data do not allow a precise link between inventive output and the intellectual effort behind its creation. Without assuming that each patent requires the same effort in the same time window before filing, we cannot disentangle the effect on output from the effect on time spent. Because implications for abatement policy vary based on the mechanism at play, it is important to disentangle them in future work. Second, our estimates may be overly conservative. Not all innovative activity is captured in patents. Many innovations are kept secret and do not show up in patent records. In addition, our back-of-the-envelope calculation obscures further spillover margins of innovation identified in the literature, which are hard to measure (Jones and Summers, 2020). Third, the cost calculation relies heavily on private value estimates of patented inventions. The fact that estimates in the literature vary widely induces considerable uncertainty about the economic magnitudes of the estimated effect. With these drawbacks in mind, we hope this paper is a useful contribution to the cost-benefit analyses underlying rational abatement policies.

CRediT authorship contribution statement

Felix Bracht: Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Dennis Verhoeven:** Writing – review & editing, Visualization, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Appendix

A.1. Additional tables and figures

See Tables 10–18.

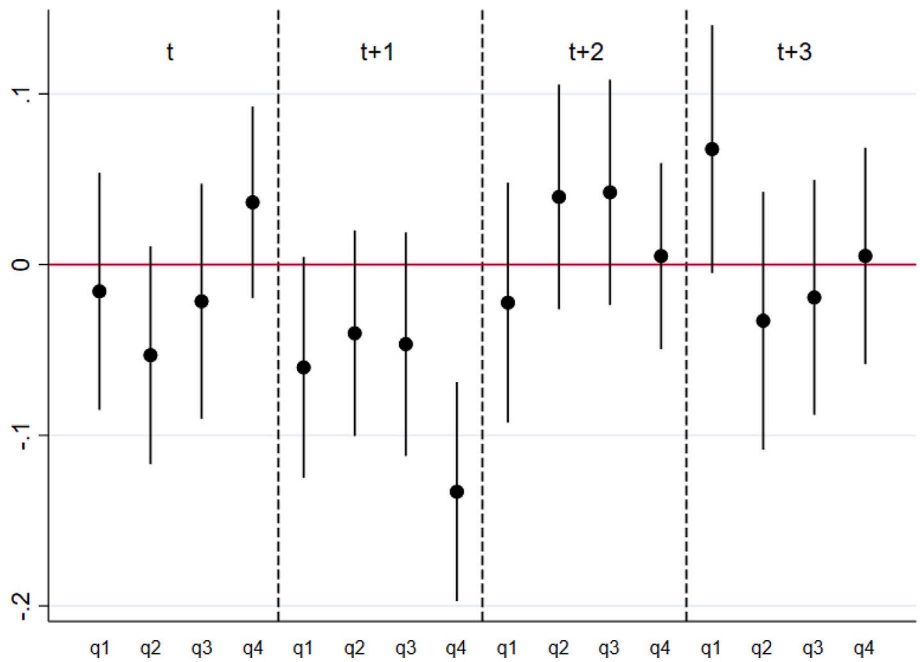

Fig. 8. Timing by quarter.
Notes: Analysis of the timing of the effect of air pollution on (patented) innovation output by quarter. The dots are the coefficients (with 95% confidence bands) from 12 regressions with the number of patented innovations in different quarters following the beginning the year in which air pollution is measured (year *t*). All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel.

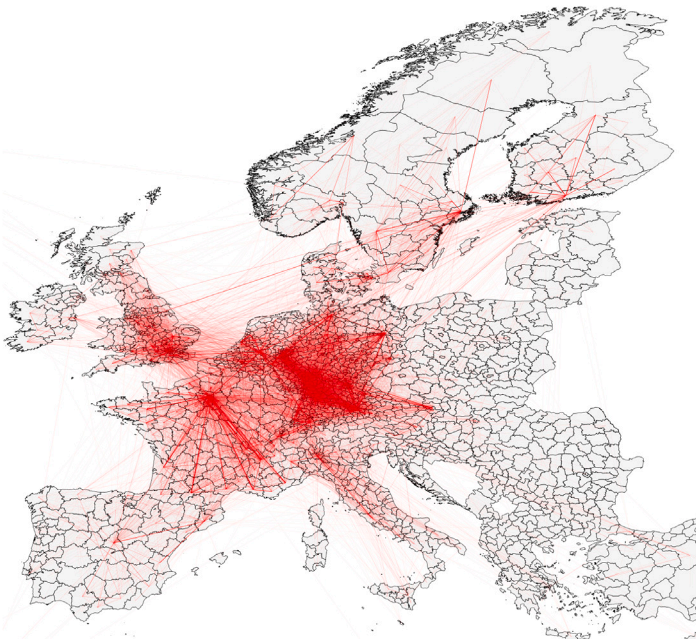

Fig. 9. Inventor relocation patterns across NUTS-3 regions.
Notes: Inventor relocation between NUTS-3 regions for the sample period 2001–2012 based on address information on consecutive patent documents. Width of line indicates relocation frequency between two regions.

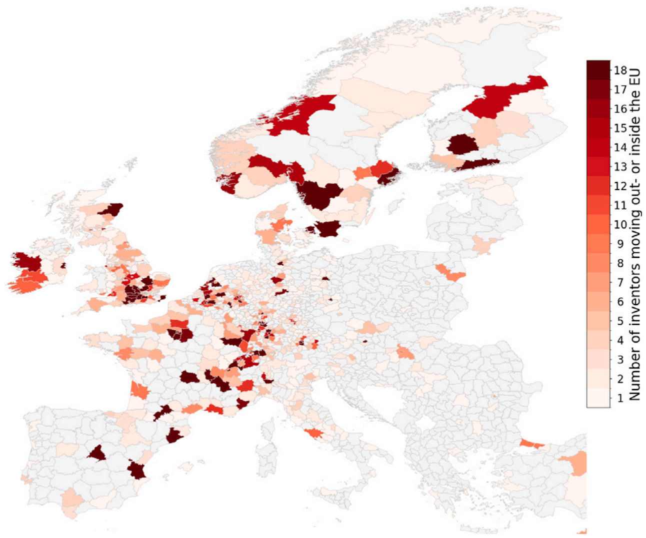

Fig. 10. Inventor relocation patterns between NUTS-3 regions and non-European regions.
Notes: Heat map of inventor relocation patterns between NUTS-3 regions and non-European regions. Color coding is based on the number of inventors that move from or to regions outside Europe.

A.2. Data cleaning inventor relocation

Morrison et al. (2017) provides a disambiguation of inventors on patents. We assign NUTS-3 regions to corresponding geocodes of inventor addresses. Next, we trace the relocation of inventors based on changes in NUTS-3 codes of patents inventors' filed in chronological order. Any address outside NUTS-3 regions is treated equally as 'outside Europe' because we are not interested in movements to areas outside Europe. In some cases, the history of filed patents stating the addresses of the individual inventor does not provide a clear relocation pattern. Specifically, the history of patents for an individual inventor shows frequent alternation between two or more NUTS-3 regions within a short period. Such patterns are unlikely to reflect true relocations of inventors. One

Table 10
Effects across the quality distribution.

	Cites				Fam. Size			
Dep. Variable	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
$PM2.5$	−0.073**	−0.12***	−0.042	−0.093**	−0.054	−0.015	−0.17***	−0.087*
	(0.031)	(0.036)	(0.034)	(0.043)	(0.037)	(0.042)	(0.037)	(0.044)
$CF\ residuals$	0.062**	0.11***	0.039	0.093**	0.043	0.0064	0.16***	0.088**
	(0.031)	(0.036)	(0.033)	(0.042)	(0.037)	(0.041)	(0.036)	(0.044)
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
KP F-stat.	✓	✓	✓	✓	✓	✓	✓	✓
Mean PM2.5	40.85	42.22	38.00	41.08	45.28	39.14	35.83	34.34
Mean Dependent	13.73	13.68	13.62	13.64	13.71	13.65	13.64	13.46
N	56.85	33.68	34.79	35.85	57.59	30.00	40.18	36.28
N	11604	11340	11232	11124	11376	11316	11268	10800

Notes: Estimates from the second-stage Poisson regression with the number of patented inventions in each of four quality quartile bins as the dependent variable. The first four columns use the total number of citations received by a patent as a quality measure. The last four columns use assess quality by how many distinct filings an invention was the subject of. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11
Placebo test: The effect of future air pollution on current innovation output.

	No IV	Inversions IV		Ventilation IV		Both IVs	
Dep. Variable:	Patents	Patents	$PM2.5$	Patents	$PM2.5$	Patents	$PM2.5$
$PM2.5$	−0.0061**	−0.036		−0.037		−0.038	
	(0.0027)	(0.031)		(0.030)		(0.025)	
$Share\ Inversions$			3.76***				3.51***
			(0.58)				(0.57)
$Share\ low\ VC$					305.9***		289.9***
					(42.4)		(42.7)
$CF\ residuals$		0.030		0.031		0.032	
		(0.032)		(0.029)		(0.025)	
NUTS-3 FE	✓	✓	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓
KP F-stat.			42.44		51.95		43.33
N	10747	10747	10747	10747	10747	10747	10747

Notes: Placebo test estimates of the effect of air pollution in $t+1$ on innovation in t . Innovation in year t is measured by the number of patent filings in the upcoming year (as in the baseline). $PM2.5$, its instruments, and the control function residual are measured in $t+1$. Variables common to the first- and second-stage equations are measured in t. The first column shows the results for the Poisson regression with no control function. The second and third columns show the second- and first-stage results of a Poisson model and a linear control function with thermal inversions as an instrument, where the first-stage residuals are controlled for in the second-stage estimation. In the fourth and fifth columns, we use ventilation capability as an instrument. The last two columns show the results when using both instruments jointly. $Share\,of\,Inversions$ measures the share of days in a given year on which a thermal inversion occurs. $Share\,of\,low\,VC$ measures, for a given year, the average number of consecutive days with a low ventilation coefficient (divided by 365 days). We define a low ventilation coefficient as any daily average ventilation coefficient that is equal or below the $20th$ percentile of the distribution of daily average ventilation coefficients within a NUTS-3 region. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Compared to the baseline, the final year dropped from the estimation because of the variables that are shifted back by 1 year. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

reason might be that inventors alternate between office and residence addresses for the patents they had filed. This issue is described in detail by Zacchia (2018) using the same dataset to trace the relocation of inventors. We follow his procedure to clean the data for unreasonable relocation patterns by assigning inventors to NUTS-3 regions in the following way:

1. We split the patent history of each inventor into temporal intervals, which are delimited by the years when a different NUTS-3 occurs for the first time.
2. To each of these intervals we assign a unique NUTS-3 region by choosing the one that is observed most frequently.
3. Ties are resolved conservatively in favor of the region that occurred earlier in the inventor's patent history.

Table 12
Timing of the effect.

Dep. Var.: Patents	t	t+1 (base)	t+2	t+3	t+4
$PM2.5$	−0.0081	−0.072***	0.030	−0.0079	0.010
	(0.025)	(0.022)	(0.025)	(0.025)	(0.030)
$CF\ residuals$	0.0026	0.068***	−0.028	0.0086	−0.0059
	(0.025)	(0.022)	(0.025)	(0.026)	(0.030)
NUTS-3 FE	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓
KP F-stat.	34.89	41.46	30.28	18.72	20.20
Mean PM2.5	13.67	13.78	13.93	13.95	14.07
Mean Dependent	118.67	123.11	125.91	128.62	130.11
N	11429	11724	10736	9760	8793

Notes: Analysis of the timing of the effect of air pollution on (patented) innovation output. The columns show the second-stage estimates for different years in which patent output is considered. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13
Aggregation of the time window.

Dep. Variable:	3 years	4 years	6 years
$PM2.5$	−0.15***	−0.20	0.14
	(0.054)	(0.18)	(0.12)
$CF\ residuals$	0.14**	0.18	−0.16
	(0.054)	(0.17)	(0.12)
NUTS-3 FE	✓	✓	✓
Country-Year FE	✓	✓	✓
Controls	✓	✓	✓
KP F-stat.	13.54	1.50	3.74
N	4512	3405	2274

Notes: Changing the granularity of the time dimension. The columns show the second-stage estimates for different aggregation levels of the time window. All estimations control for NUTS-3 fixed effects, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14
Different thresholds for a low VC.

Dep. Variable:	p10	p15	p20 (base)	p25
$PM2.5$	−0.060**	−0.063**	−0.072***	−0.058***
	(0.029)	(0.025)	(0.022)	(0.019)
$CF\ residuals$	0.056*	0.059**	0.068***	0.054***
	(0.029)	(0.025)	(0.022)	(0.018)
NUTS-3 FE	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
KP F-stat.	28.46	32.87	41.46	51.26
N	11724	11724	11724	11724

Notes: Results of regressions with alternative definition of the ventilation capability instrument. The columns show the second-stage estimates for different definitions for the ventilation capability threshold. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15
Weekday versus weekend inversions.

	Inversions only			Both IVs		
TI version:	Base	Weekend	Weekday	Base	Weekend	Weekday
$PM2.5$	−0.048	0.0018	−0.081*	−0.072***	−0.050**	−0.095***
	(0.030)	(0.032)	(0.044)	(0.022)	(0.023)	(0.026)
$CF\ residuals$	0.044	−0.0061	0.077*	0.068***	0.046**	0.092***
	(0.030)	(0.031)	(0.044)	(0.022)	(0.023)	(0.026)
NUTS-3 FE	✓	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
KP F-stat.	41.46	29.54	35.08	41.46	29.54	35.08
N	11724	11724	11724	11724	11724	11724

Notes: Results of regressions with alternative definitions of the thermal inversion instrument. The first column shows the second-stage estimates for the baseline model. The second column model uses only thermal inversions on weekend days. The third column uses only thermal inversions on weekdays. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16
Regression weights.

Regression weights:	None	Population	Patents
$PM2.5$	−0.072***	−0.071***	−0.068***
	(0.022)	(0.022)	(0.024)
$CF\ residuals$	0.068***	0.067***	0.064***
	(0.022)	(0.022)	(0.025)
NUTS-3 FE	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓
Country-Year FE	✓	✓	✓
Controls	✓	✓	✓
KP F-stat.	41.46	41.80	33.93
N	11724	11724	10668

Notes: Results of regressions with weights. The first column shows the second-stage estimates for the baseline model. The second column weights for the natural logarithm of the population in the first year of the sample. The third column weights for the natural logarithm of the patent count in the first year of the sample. For this model, regions without patents in the first year of the sample receive no weight and are dropped. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17
Dropping outliers in terms of patenting and air pollution.

	Patents			$PM2.5$		
Drop outliers:	<1pct.	>99pct.	both	<1pct.	>99pct.	both
$PM2.5$	−0.071***	−0.079***	−0.079***	−0.074***	−0.069***	−0.071***
	(0.022)	(0.021)	(0.021)	(0.022)	(0.023)	(0.023)
CF *residuals*	0.068***	0.074***	0.074***	0.070***	0.066***	0.068***
	(0.022)	(0.020)	(0.020)	(0.022)	(0.023)	(0.023)
NUTS-3 FE	✓	✓	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
KP F-stat.	40.85	41.04	40.44	41.40	39.34	39.30
N	11628	11616	11520	11616	11616	11508

Notes: Results of regressions where outliers are dropped. The first set of three columns drops outlier regions in terms of the total number of patents in our time frame (respectively, regions below the 1st, above the 99th, and both are dropped). The second set of three columns shows the analogous result for the average level of air pollution. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18
Effect on inventor migration — high confidence migration.

	Emigration		Immigration	
Dep. Var.	Inventors	Inv. patents	Inventors	Inv. patents
$PM2.5$	−0.090	−0.20	−0.015	−0.081
	(0.13)	(0.19)	(0.10)	(0.16)
CF *residuals*	0.085	0.20	0.020	0.082
	(0.13)	(0.19)	(0.11)	(0.16)
NUTS-3 FE	✓	✓	✓	✓
NUTS-3 Slopes	✓	✓	✓	✓
Country-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
KP F-stat.	33.27	33.27	36.05	36.05
Mean PM2.5	13.09	13.09	13.13	13.13
Mean Dep. Var.	1.58	7.94	1.63	7.78
N	8460	8460	8868	8868

Notes: Second-stage results of the analyses of the effect on inventor migration, where we restrict to inventor moves that could be determined with high confidence. *Inventors* refers to the number of inventors that were identified as moving between regions. *Inv. patents* weights the number of migrating inventors by the number of patents inventors filed within 5 years before the relocation. The first two columns refer to inventors migrating out of the focal region, the next two columns refer to inventors migrating into the focal region. The final column estimates the effect on the number of patents in a region divided by the region's population. All estimations control for fixed effects and linear time trends at the NUTS-3 level, country-by-year fixed effects, population (in logs), and weather conditions (first- and second-order terms of average wind speed, humidity, temperature, and precipitation). The KP F-stat. refers to the Kleibergen–Paap F-statistic to test for instrument strength (Kleibergen and Paap, 2006; Stock and Yogo, 2002). Estimations are based on the largest possible fully balanced panel. Standard errors are reported in parentheses and clustered at the NUTS-3 level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Aghion, P., Akcigit, U., Hyytinen, A., Toivanen, O., 2017. The Social Origins of Inventors. NBER Working Paper No. 24110, National Bureau of Economic Research.

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., Van Reenen, J., 2016. Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. J. Polit. Econ. 124 (1), 1–51.

Aghion, P., Howitt, P., 1990. A Model of Growth Through Creative Destruction. NBER Working Paper No. 3223, National Bureau of Economic Research.

Aguilar-Gomez, S., Dwyer, H., Graff Zivin, J., Neidell, M., 2022. This is air: The "nonhealth" effects of air pollution. Annu. Rev. Resour. Econ. 14, 403–425.

Anderson, M.L., 2020. As the wind blows: The effects of long-term exposure to air pollution on mortality. J. Eur. Econom. Assoc. 18 (4), 1886–1927.

Aragón, F.M., Miranda, J.J., Oliva, P., 2017. Particulate matter and labor supply: The role of caregiving and non-linearities. J. Environ. Econ. Manag. 86, 295–309.

Arceo, E., Hanna, R., Oliva, P., 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico city. Econ. J. 126 (591), 257–280.

Archsmith, J., Heyes, A., Saberian, S., 2018. Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. J. Assoc. Environ. Resour. Econ. 5 (4), 827–863.

Arque-Castells, P., Spulber, D.F., 2022. Measuring the private and social returns to R&D: Unintended spillovers versus technology markets. J. Polit. Econ. 130 (7), 1860–1918.

Avila Uribe, A., 2023. The Effect of Air Pollution on US Aggregate Production. Technical Report 39, Department of Geography and Environment, LSE.

Barreca, A.I., Neidell, M., Sanders, N.J., 2021. Long-run pollution exposure and mortality: Evidence from the acid rain program. J. Public Econ. 200, 104440.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., Van Reenen, J., 2019. Who becomes an inventor in America? The importance of exposure to innovation. Q. J. Econ. 134 (2), 647–713.

Bessen, J., 2009. Estimates of patent rents from firm market value. Res. Policy 38 (10), 1604–1616.

Bloom, N., Schankerman, M., Reenen, J.V., 2013. Identifying technology spillovers and product market rivalry. Econometrica 81 (4), 1347–1393.

Bovenberg, A.L., Smulders, S., 1995. Environmental quality and pollution-augmenting technological change in a two-sector endogenous growth model. J. Public Econ. 57 (3), 369–391.

Bovenberg, A.L., Smulders, S.A., 1996. Transitional impacts of environmental policy in an endogenous growth model. Internat. Econom. Rev. 861–893.

Bronzini, R., Piselli, P., 2016. The impact of R&D subsidies on firm innovation. Res. Policy 45 (2), 442–457.

Chang, T., Zivin, J.G., Gross, T., Neidell, M., 2016. Particulate pollution and the productivity of pear packers. Am. Econ. J. Econ. Policy 8 (3), 141–169.

Chang, T.Y., Zivin, J.G., Gross, T., Neidell, M., 2019. The effect of pollution on worker productivity: Evidence from call center workers in China. Am. Econ. J. Appl. Econ. 11 (1), 151–172.

Chen, S., Oliva, P., Zhang, P., 2022. The effect of air pollution on migration: Evidence from China. J. Dev. Econ. 156, 102833.

Columbia University, C.f.I.E.S.I.N.C., 2018. Gridded population of the world, version 4 (GPWv4): Population count, revision 11.

Cook, N., Heyes, A., Rivers, N., 2023. Clean air and cognitive productivity: Effect and adaptation. J. Assoc. Environ. Resour. Econ. 10 (5), 1265–1308.

de Rassenfosse, G., Kozak, J., Seliger, F., 2019. Geocoding of worldwide patent data. Sci. Data 6 (1), 260.

Dechezleprêtre, A., Rivers, N., Stadler, B., 2019. The Economic Cost of Air Pollution: Evidence from Europe. OECD Economics Department Working Papers No. 1584, OECD.

Deng, Y., 2007. Private value of European patents. Eur. Econ. Rev. 51 (7), 1785–1812.

Deryugina, T., Heutel, G., Miller, N.H., Molitor, D., Reif, J., 2019. The mortality and medical costs of air pollution: Evidence from changes in wind direction. Amer. Econ. Rev. 109 (12), 4178–4219.

Dong, R., Fisman, R., Wang, Y., Xu, N., 2021. Air pollution, affect, and forecasting bias: Evidence from Chinese financial analysts. J. Financ. Econ. 139 (3), 971–984.

Duque, V., Gilraine, M., 2022. Coal use, air pollution, and student performance. J. Public Econ. 213, 104712.

Ebenstein, A., Lavy, V., Roth, S., 2016. The long-run economic consequences of high- stakes examinations: Evidence from transitory variation in pollution. Am. Econ. J. Appl. Econ. 8 (4), 36–65.

European Commission, 2020. Science, research and innovation performance of the EU 2020.

European Environment Agency, 2020. Air quality in Europe — 2020 report.

European Parliament and Council of the European Union, 2008. Directive 2008/50/EC. OJ L 152, 1–44.

Fischer, T., Leidinger, J., 2014. Testing patent value indicators on directly observed patent value — An empirical analysis of ocean tomo patent auctions. Res. Policy 43 (3), 519–529.

Fu, S., Viard, V.B., Zhang, P., 2021. Air pollution and manufacturing firm productivity: Nationwide estimates for China. Econ. J. 131 (640), 3241–3273.

Gambardella, A., Harhoff, D., Verspagen, B., 2008. The value of European patents. Eur. Manag. Rev. 5 (2), 69–84.

Garratt, J.R., 1994. The atmospheric boundary layer. Earth-Sci. Rev. 37 (1–2), 89–134.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., et al., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30 (14), 5419–5454.

Genc, D.D., Yesilyurt, C., Tuncel, G., 2010. Air pollution forecasting in ankara, Turkey using air pollution index and its relation to assimilative capacity of the atmosphere. Environ. Monit. Assess. 166 (1), 11–27.

Gramsch, E., Cáceres, D., Oyola, P., Reyes, F., Vásquez, Y., Rubio, M.A., Sánchez, G., 2014. Influence of surface and subsidence thermal inversion on PM2. 5 and black carbon concentration. Atmos. Environ. 98, 290–298.

Griliches, Z., 1990. Patent statistics as economic indicators: A survey. J. Econ. Lit. 28 (4), 1661.

Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Papers No. 8498.

Hall, B.H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. Rand J. Econ. 16–38.

Hammer, M.S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A.M., Hsu, N.C., Levy, R.C., Garay, M.J., Kalashnikova, O.V., Kahn, R.A., et al., 2020. Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). Environ. Sci. Technol. 54 (13), 7879–7890.

Hanna, R., Oliva, P., 2015. The effect of pollution on labor supply: Evidence from a natural experiment in Mexico city. J. Public Econ. 122, 68–79.

Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation frequency and the value of patented inventions. Rev. Econ. Stat. 81 (3), 511–515.

Harhoff, D., Scherer, F.M., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. Res. Policy 32 (8), 1343–1363.

He, J., Liu, H., Salvo, A., 2019. Severe air pollution and labor productivity: Evidence from industrial towns in China. Am. Econ. J. Appl. Econ. 11 (1), 173–201.

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

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049.

Holzworth, G.C., 1967. Mixing depths, wind speeds and air pollution potential for selected locations in the United States. J. Appl. Meteorol. 6 (6), 1039–1044.

Holzworth, G.C., 1972. Mixing Heights, Wind Speeds, and Potential for Urban Air Pollution Throughout the Contiguous United States. US Government Printing Office.

Hou, X., Fei, D., Kang, H., Zhang, Y., Gao, J., 2018. Seasonal statistical analysis of the impact of meteorological factors on fine particle pollution in China in 2013–2017. Nat. Hazards 93 (2), 677–698.

Isen, A., Rossin-Slater, M., Walker, W.R., 2017. Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970. J. Polit. Econ. 125 (3), 848–902.
Jones, B., Summers, L., 2020. A Calculation of the Social Returns to Innovation. NBER Working Paper No. 27863, National Bureau of Economic Research, Cambridge, MA.
Kaltenberg, M., Jaffe, A., Lachman, M.E., 2021. Matched inventor ages from patents, based on web scraped sources.
Khanna, G., Liang, W., Mobarak, A.M., Song, R., 2021. The Productivity Consequences of Pollution-Induced Migration in China. NBER Working Papers No. 28401, National Bureau of Economic Research.
Kleibergen, F., Paap, R., 2006. Generalized reduced rank tests using the singular value decomposition. J. Econometrics 133 (1), 97–126.
Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. Q. J. Econ. 132 (2), 665–712.
Koolen, C.D., Rothenberg, G., 2019. Air pollution in Europe. ChemSusChem 12 (1), 164.
Künn, S., Palacios, J., Pestel, N., 2023. Indoor air quality and strategic decision making. Manage. Sci..
La Nauze, A., Severnini, E.R., 2021. Air Pollution and Adult Cognition: Evidence from Brain Training. NBER Working Paper No. 28785.
Landrigan, P.J., Fuller, R., Acosta, N.J., Adeyi, O., Arnold, R., Basu, N.N., Baldé, A.B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A., Murray, C.J., Ndahimananjara, J.D., Perera, F., Potočnik, J., Preker, A.S., Ramesh, J., Rockström, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W.A., van Schayck, O.C., Yadama, G.N., Yumkella, K., Zhong, M., 2018. The lancet commission on pollution and health.
Lichter, A., Pestel, N., Sommer, E., 2017. Productivity effects of air pollution: Evidence from professional soccer. Labour Econ. 48, 54–66.
Lin, W., Wooldridge, J.M., 2019. Testing and correcting for endogeneity in nonlinear unobserved effects models. In: Panel Data Econometrics. Elsevier, pp. 21–43.
Morrison, G., Riccaboni, M., Pammolli, F., 2017. Disambiguation of patent inventors and assignees using high-resolution geolocation data. Sci. Data 4 (1), 1–21.
Muñoz-Sabater, J., 2019. ERA5-land hourly data from 2001 to present.
Myers, K.R., Lanahan, L., 2022. Estimating spillovers from publicly funded R&D: Evidence from the US department of energy. Amer. Econ. Rev. 112 (7), 2393–2423.
Pakes, A., 1986. Patents as options: Some estimates of the value of holding European patent stocks. Econometrica 54 (4), 755.
Power, M.C., Adar, S.D., Yanosky, J.D., Weuve, J., 2016. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. Neurotoxicology 56, 235–253.
Ravenscraft, D., Scherer, F.M., 1982. The lag structure of returns to research and development. Appl. Econ. 14, 603–620.
Romer, P.M., 1990. Endogenous technological change. J. Political Econ. 98 (5, Part 2), S71–S102.
Schankerman, M., Pakes, A., 1986. Estimates of the value of patent rights in European countries during the post-1950 period. Econ. J. 96 (384), 1052–1076.
Scherer, F.M., 1965. Corporate inventive output, profits, and growth. J. Polit. Econ. 73 (3), 290–297.
Schmoch, U., 2008. Concept of a Technology Classification for Coun-try Comparisons 1 Final Report to the World Intellectual Property Organisation (WIPO). Technical Report.
Schmookler, J., 1966. Invention and Economic Growth. Harvard University Press.
Stafford, T.M., 2015. Indoor air quality and academic performance. J. Environ. Econ. Manag. 70, 34–50.
Stevenson, A., Nardotto, M., Toivanen, O., Takalo, T., 2022. Essays on the economics of patent rights.
Stock, J., Yogo, M., 2002. Testing for Weak Instruments in Linear IV Regression. NBER Working Paper No. 0284, National Bureau of Economic Research, Inc.
Stull, R.B., 1988. An Introduction to Boundary Layer Meteorology, vol. 13. Springer Science & Business Media.
Torssander, J., Erikson, R., 2009. Stratification and mortality—A comparison of education, class, status, and income. Eur. Sociol. Rev. 26 (4), 465–474.
Trajtenberg, M., 1990. A penny for your quotes: Patent citations and the value of innovations. RAND J. Econ. 172–187.
Underwood, E., 2017. The polluted brain. Science 355 (6323), 342–345.
US EPA (U.S. Environmental Protection Agency), 1997. 40 CFR part 50. National ambient air quality standards for particulate matter. Final rule. Fed Reg 62, pages 38652–38760.
Vallero, D., 2014. Chapter 3 - the science of air pollution. In: Vallero, D. (Ed.), Fundamentals of Air Pollution (Fifth Edition), Fifth ed. Academic Press, Boston, pp. 43–81.
Van Donkelaar, A., Martin, R.V., Li, C., Burnett, R.T., 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. 53 (5), 2595–2611.
WHO, 2016. Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease. World Health Organization, p. 121.
Wooldridge, J.M., 1999. Distribution-free estimation of some nonlinear panel data models. J. Econometrics 90 (1), 77–97.
Xu, G., Jiao, L., Zhang, B., Zhao, S., Yuan, M., Gu, Y., Liu, J., Tang, X., et al., 2017. Spatial and temporal variability of the PM2. 5/PM10 ratio in Wuhan, central China. Aerosol Air Qual. Res. 17 (3), 741–751.
Zacchia, P., 2018. Benefiting colleagues but not the city: Localized effects from the relocation of superstar inventors. Res. Policy 47 (5), 992–1005.
Zacchia, P., 2020. Knowledge spillovers through networks of scientists. Rev. Econ. Stud. 87 (4), 1989–2018.
Zivin, J.G., Neidell, M., 2012. The impact of pollution on worker productivity. Amer. Econ. Rev. 102 (7), 3652–3673.
Zivin, J.G., Neidell, M., 2018. Air pollution's hidden impacts. Science 359 (6371), 39–40.